

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 6, 2003, 14:32:48 ; Search time 6517 Seconds

(without alignments)
11136.056 Million cell updates/sec

Title: us-09-854-300-7

Perfect score: 1774

Sequence: 1 agctgagctccacgcgcgt.....atagcgcaagttctctctcag 1774

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapept 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBank1:*

1: gb Da:*

2: gb Htg:*

3: gb In:*

4: gb Om:*

5: gb Ov:*

6: gb Pat:*

7: gb Ph:*

8: gb Pl:*

9: gb Pr:*

10: gb Ro:*

11: gb Sts:*

12: gb Sy:*

13: gb Un:*

14: gb Vi:*

15: em Ba:*

16: em Fun:*

17: em Hum:*

18: em In:*

19: em Mu:*

20: em Om:*

21: em Or:*

22: em Ov:*

23: em Pat:*

24: em Ph:*

25: em Pl:*

26: em Ro:*

27: em Sts:*

28: em Un:*

29: em Vi:*

30: em Htg Hum:*

31: em Htg Inv:*

32: em Htg Other:*

33: em Htg Mus:*

34: em Htg Pln:*

35: em Htg Rod:*

36: em Htg Mam:*

37: em Htg Vtl:*

38: em Sy:*

39: em Htgo Hum:*

40: em Htgo Mus:*

41: em Htgo Other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1662.8	93.7	2796	6 AX337968	AX337968 Sequence
2	1659.6	93.6	2830	6 AK074264	AK074264 Homo sapi
3	1431.4	80.7	2876	6 AX576252	AX576252 Sequence
4	1280.8	72.2	1297	6 AF394689	AF394689 Homo sapi
5	1205.2	67.9	1928	10 BC010477	BC010477 Mus muscu
6	1203.6	67.8	1936	10 AB041548	AB041548 Mus muscu
7	1137.2	64.1	1513	10 AY112656	AY112656 Mus muscu
8	1106.8	62.4	1287	10 AF426411	AF426411 Mus muscu
9	714.4	40.3	1871	9 AK027169	AK027169 Homo sapi
10	674.4	38.0	169812	2 AL391315	AL391315 Human DNA
11	437.4	38.0	207775	2 AL591467	AL591467 Homo sapi
12	432.2	24.2	252305	2 AC114183	AC114183 Rattus no
13	428.2	23.6	239385	2 EX470094	EX470094 Mus muscu
14	418.2	23.6	239385	2 AC109686	AC109686 Rattus no
15	392.2	22.1	404	6 AX071980	AX071980 Sequence
16	372.6	21.0	646	6 AX341568	AX341568 Sequence
17	310	17.5	435	6 AX396961	AX396961 Sequence
18	280.4	15.8	1020	10 AY155441	AY155441 Mus muscu
19	280.4	15.8	195962	2 AC115936	AC115936 Mus muscu
20	278.4	15.7	244434	2 AC111461	AC111461 Rattus no
21	268.8	15.2	1391	9 AB070023	AB070023 Macaca fa
22	268.8	15.2	1395	9 AB070058	AB070058 Macaca fa
23	266.8	15.0	1429	9 AK098524	AK098524 Homo sapi
24	266.8	15.0	1430	9 BC022038	BC022038 Homo sapi
25	266.8	15.0	1446	9 AF447589	AF447589 Homo sapi
26	266.8	15.0	120984	9 AC006463	AC006463 Homo sapi
27	257.8	14.5	1148	10 AY155440	AY155440 Mus muscu
28	256.4	14.5	1328	9 BC023264	BC023264 Homo sapi
29	254.8	14.4	1355	6 AR306585	AR306585 Sequence
30	254.8	14.4	1355	6 AX061659	AX061659 Sequence
31	253.2	14.3	1261	9 AK098654	AK098654 Homo sapi
32	220	12.4	119995	10 AL731648	AL731648 Mouse DNA
33	220	12.4	252305	2 EX470094	EX470094 Mus muscu
34	217.4	12.3	239385	2 AC109686	AC109686 Rattus no
35	209	11.8	1200	6 AX083416	AX083416 Sequence
36	209	11.8	1527	6 AX083426	AX083426 Sequence
37	209	11.8	2001	9 BC019355	BC019355 Homo sapi
38	209	11.8	2005	6 BD127528	BD127528 Homo sapi
39	209	11.8	2005	9 AK075141	AK075141 Homo sapi
40	209	11.8	2021	9 BC045743	BC045743 Homo sapi
41	209	11.8	2063	6 BD063226	BD063226 Secreterd
42	209	11.8	2452	6 BD127402	BD127402 Primer fo
43	209	11.8	2452	9 AK074985	AK074985 Homo sapi
44	207.4	11.7	2013	9 BC032328	BC032328 Homo sapi
45	201.2	11.3	2328	6 AX644275	AX644275 Sequence

ALIGNMENTS

RESULT 1

AX337968

LOCUS AX337968 2796 bp DNA linear PAT 09-JAN-2002

DEFINITION Sequence 9 from Patent WO0194391.

ACCESSION AX337968

VERSION AX337968.1 GI:18128679

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 Yue H., He A., Nguyen D.B., Yao M.G., Bandman O., Burford N., Tang Y.T., Xu Y., Hafalia A., Azimaei Y. and Walla N.K. Intracellular signaling proteins

JOURNAL Patent: WO 0194391-A 9 13-DEC-2001;
 Incyte Genomics, Inc. (US)
 Location/Qualifiers
 FEATURES
 source
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /note="Incyte ID No: 279168CE1"
 BASE COUNT 821 a 548 c 620 g 807 t
 ORIGIN
 Query Match 93.7%; Score 1662.8; DB 6; Length 2796;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 1690; Conservative 0; Mismatches 8; Indels 3; Gaps 2;
 QY 74 CCGAGAGCTGATCTCTGGGCAACTGTGTGTGAAGCTTGGCTCTGGCTCGAG 133
 DB 28 CCGAGAGCTGATCTCTGGGCAACTGTGTGTGAAGCTTGGCTCTGGCTCGAG 87
 QY 134 TAGCTGAGCTCCCGAGTCACTCACTCTCCCACTTGGCGGCACTTCTGAG 193
 DB 88 TAGCTGAGCTCCCGAGTCACTCACTCTCCCACTTGGCGGCACTTCTGAG 147
 QY 194 ACCAGGCTCTGCGAAGCGCTAGAGAGCGGCGTCAAGGCGCTAGAGGAACTCGGAGC 253
 DB 148 ACCAGGCTCTGCGAAGCGCTAGAGAGCGGCGTCAAGGCGCTAGAGGAACTCGGAGC 207
 QY 254 GCGCGGCGCATGGGCGCGCGCTGGGCGCGGCGCTGGGCGCGGCGCTGGGCGCTGG 313
 DB 208 GCGCGGCGCATGGGCGCGCGCTGGGCGCGGCGCTGGGCGCGGCGCTGGGCGCTGG 267
 QY 314 TCCAGATTGCTGCAAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 373
 DB 268 TCCAGATTGCTGCAAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 327
 QY 374 GGGGCTGAGAGAGTGTGGAACCGGCTCAACGCTGCTGAGCGGCTTCCGACACGAG 433
 DB 328 GGGGCTGAGAGAGTGTGGAACCGGCTCAACGCTGCTGAGCGGCTTCCGACACGAG 387
 QY 434 GTGAACTGATGCTGTGGAAGCTGAGCGAGAGAGCGGCTGAGCGGCTGAGCGGCTG 493
 DB 388 GTGAACTGATGCTGTGGAAGCTGAGCGAGAGAGCGGCTGAGCGGCTGAGCGGCTG 447
 QY 494 GAGCGTGTGCTGGGCGCTCTGTGATACCGCGCGCGCGCGCGCTTGAAGCGCTGAG 553
 DB 448 GAGCGTGTGCTGGGCGCTCTGTGATACCGCGCGCGCGCGCGCTTGAAGCGCTGAG 507
 QY 554 CCGCACAAGATTTCACGCTGCCACGCTTGGGAGAGCAACGCTGAGAGCTCTTGGTGG 613
 DB 508 CCGCACAAGATTTCACGCTGCCACGCTTGGGAGAGCAACGCTGAGAGCTCTTGGTGG 567
 QY 614 GCGCTGATCAACGCGCGGCGGCGCTGACCTTGGCAGACAAGATTCATGCTTATGAG 673
 DB 568 GCGCTGATCAACGCGCGGCGGCGCTGACCTTGGCAGACAAGATTCATGCTTATGAG 627
 QY 674 AGATGGCGCTGTGAGCGCTGATCTTTAACTTCCCGGAGACCGGCAATGAGTCAATCCC 733
 DB 628 AGATGGCGCTGTGAGCGCTGATCTTTAACTTCCCGGAGACCGGCAATGAGTCAATCCC 687
 QY 734 ATGCTACCGCGGCGGCGAGTACATGTTGGAATCAATGATCGGCAATCGAAAGGACA 793
 DB 688 ATGCTACCGCGGCGGCGAGTACATGTTGGAATCAATGATCGGCAATCGAAAGGACA 747
 QY 794 AAATTTCTGCAATCTATTCAAGAGGATACAGTACATGCTGATAGAGTAAAGTAAAG 853
 DB 748 AAATTTCTGCAATCTATTCAAGAGGATACAGTACATGCTGATAGAGTAAAGTAAAG 807
 QY 854 AAACATGGCGCTGTGGTGAATCAATCTATCTATTTTCTGCTTCTGCTTCTTATTT 913
 DB 808 AAACATGGCGCTGTGGTGAATCAATCTATCTATTTTCTGCTTCTGCTTCTTATTT 867
 QY 914 ATTACGCGCACTGTGGGCTATTATCTTTATTTCTGCTGAGAGGCTAGAGATGCA 973

DB 868 ATTACGCGCACTGTGGGCTATTATCTTTATTTCTGCTGAGAGGCTACGGAATGCA 927
 QY 974 AAGGCTCAAG 1033
 DB 928 AAGGCTCAAG 987
 QY 1034 CTTCATCTAG 1093
 DB 988 CTTCATCTAG 1047
 QY 1094 GCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1153
 DB 1048 GCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1107
 QY 1154 ATTTTCAATGAG 1213
 DB 1108 ATTTTCAATGAG 1167
 QY 1214 AATGTGACATGATCAAGAGCTTGGGAAATGAGGAGAGAGAGAGAGAGAGAGAG 1273
 DB 1168 AATGTGACATGATCAAGAGCTTGGGAAATGAGGAGAGAGAGAGAGAGAGAGAG 1227
 QY 1274 TTCAAGTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1333
 DB 1228 TTCAAGTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1287
 QY 1334 CGCAGGAG 1393
 DB 1288 CGCAGGAG 1347
 QY 1394 GAGGAG 1453
 DB 1348 GAGGAG 1407
 QY 1454 GTGGCAG 1513
 DB 1408 GTGGCAG 1467
 QY 1514 AATCAAG 1573
 DB 1468 AATCAAG 1527
 QY 1574 ACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1633
 DB 1528 ACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1587
 QY 1634 TTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1693
 DB 1588 TTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1647
 QY 1694 AATGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1753
 DB 1648 A--TGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1704
 QY 1754 CATAGGCAAGTTCTCTCAG 1774
 DB 1705 CATAGGCAAGTTCTCTCAG 1725

RESULT 2
 AK074264 2830 bp mRNA linear PRI 15-FEB-2002
 LOCUS AK074264
 DEFINITION Homo sapiens cDNA FLJ33684 f1s, clone HEP09821.
 ACCESSION AK074264
 VERSION AK074264.1 GI:18676818
 KEYWORDS oligo capding; f1s (full insert sequence).
 SOURCE
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 Kawabata, A., Hiki, J. T., Kobatake, N., Inagaki, H., Ikema, Y.,
 Okamoto, S., Ohtani, R., Ota, T., Suzuki, Y., Odayashi, M., Nishi, T.,

TITLE
Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.
NEDO human cDNA sequencing project
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 2830)
Sugano, S., Suzuki, Y., Oka, T., Obayashi, M., Nishi, T., Isogai, T.,
AUTHORS
Shibahara, T., Tanaka, T. and Nakamura, Y.
TITLE
Direct Submission
JOURNAL
Submitted (14-PEB-2002) Sunio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure, Human Genome
Center, Shirokane-dai, 4-6-1, Minato-Ku, Tokyo 108-8639, Japan
(E-mail: cshale@ims.u-tokyo.ac.jp, Tel: 81-3-5449-5286,
Fax: 81-3-5449-5416)

COMMENT
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan: cDNA full insert sequencing;
Research Association for Biotechnology: cDNA library construction;
5'- & 3'-end one pass sequencing; Department of Virology and Human
Genome Center, Institute of Medical Science, University of Tokyo
(partly supported by Science and Technology Agency).
FEATURES
source
Location/Qualifiers
1..2830
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HEP209821"
/cell_line="Hepl2"
/cell_type="hepatoma"
/clone_1ib="HEP"
/note="Cloning vector: pME18SFL3"
251..1537
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB85033.1"
/db_xref="GI:18676819"
/translation="MGPPGAGVSCRCGCGFRLTAMCTTALSPQAPSGAEAVWT
AYLAWSRVPHGVNRTWELSEGVYGDSPLEAVAVLPPPGGALANACPTNF
TVPVMSGVQVSMALIORGGCTEADKHLAYRGAAGAVINPPGPRNVPMISH
EGAVDIYALIMLKQTKLIOSIORGIQVMTYKKGHWNVHSIFFVSPFII
TAATVGFIFYARLRNRAKROKOLKADAKKALIRLRLTLOKQKEIFPGDGS
CAVCIPLVKNLIVRLITCNHIFPKTCVDPMLERCTKMDIKALGIVDVEDG
SVSLQTVSVETSNASGSHEDNNGSTASGTAISGCTDPEPPEHVGSTNLSLVN
HEANSVAVDVIPIVDNPFEEEDTNPQETNAEIKS"

CDS

BASE COUNT 827 a 564 c 630 g 809 t
ORIGIN

Query Match 93.6%; Score 1659.6; DB 9; Length 2830;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1688; Conservative 0; Mismatches 10; Indels 3; Gaps 2;

74 CCGAGAGCTGCATCTGCGGCAACCTGTGCTGAAGCTGCTCTCTGAGCG 133
62 CCGAGAGCTGCATCTGCGGCAACCTGTGCTGAAGCTGCTCTCTGAGCG 121
134 TAGCTGCGAGCTCCCGAGTCTACCTCCCTCCCGAGCGCGCACTGCTCAAG 193
122 TAGCTGCGAGCTCCCGAGTCTACCTCCCTCCCGAGCGCGCACTGCTCAAG 181
194 ACCAGAGTCTCCCAAGCGCTAGAGAGCGCGCTGCGAGCGCGCACTGCTCAAG 253
182 ACCAGAGTCTCCCAAGCGCTAGAGAGCGCGCTGCGAGCGCGCACTGCTCAAG 241
254 GCGCGCGCATAGGAGCGCGCTGCGAGCGCGCTGCGAGCGCGCACTGCTCAAG 313
242 GCGCGCGCATAGGAGCGCGCTGCGAGCGCGCTGCGAGCGCGCACTGCTCAAG 301
314 TCCAGATTGCTGCGAGTCTCTCTGCTGAGCTGAGTCCGAGAGCACTCGGTTCCCG 373
302 TCCAGATTGCTGCGAGTCTCTCTGCTGAGCTGAGTCCGAGAGCACTCGGTTCCCG 361
374 GGGGCTGAAGCACTGTGGAACCGGTAACCAAGCTGCGAGCGGTTCCGCAACAGGGA 433
362 GGGGCTGAAGCACTGTGGAACCGGTAACCAAGCTGCGAGCGGTTCCGCAACAGGGA 421
434 GTGAACCGTAAGTGTGGAAGCTGAAGCGAGAGGAGCGCTGTACCGCCAGACTCGCGCTG 493

422 GTGAACCGTAAGTGTGGAAGCTGAAGCGAGAGGAGCGCTGTACCGCCAGACTCGCGCTG 481
494 GAGCTGTGCTGAGGCTCTGTTCCGCGCGAGCGCGCGCGCTTAAACGCTGTAC 553
482 GAGCTGTGCTGAGGCTCTGTTCCGCGCGAGCGCGCGCGCTTAAACGCTGTAC 541
554 CCGAGACGAAATTTCACGCTGCCAGGCTTGGAGAGAGCAACGTCGAGCTCTGTTG 613
542 CCGAGACGAAATTTCACGCTGCCAGGCTTGGAGAGAGCAACGTCGAGCTCTGTTG 601
614 GCCCTATCCACGCGCGCGCGCGCTGCACTTGCAGACAGATTCATCGCTTATGAG 673
602 GCCCTATCCACGCGCGCGCGCGCTGCACTTGCAGACAGATTCATCGCTTATGAG 661
674 AGATGAGGCTGTGAGACGCTGATCTTAACTTCCCGGAGCCCGGATAGGATCAACCC 733
662 AGATGAGGCTGTGAGACGCTGATCTTAACTTCCCGGAGCCCGGATAGGATCAACCC 721
734 ATGTCTACCCGAGTGCAGTGAACAATTGTCATCATGATCGCATCTGAAGGACACA 793
722 ATGTCTACCCGAGTGCAGTGAACAATTGTCATCATGATCGCATCTGAAGGACACA 781
794 AAATTCGGAATCTATTCAAAGAGGATCAAGTGAACAATGTCATGAAGTGAAGAAA 853
782 AAATTCGGAATCTATTCAAAGAGGATCAAGTGAACAATGTCATGAAGTGAAGAAA 841
854 AAACATGAGCCCTTGGTGAATCACTTCAATTTTGGTTTGTGCTCTTTTAT 913
842 AAACATGAGCCCTTGGTGAATCACTTCAATTTTGGTTTGTGCTCTTTTAT 901
914 ATTCAGCGGCAACCTGCGGCTATTTATCTTTATCTGCGCGAAGCAAGCAATGCA 973
902 ATTCAGCGGCAACCTGCGGCTATTTATCTTTATCTGCGCGAAGCAAGCAATGCA 961
974 AGAGCTCAAGCGAGAGCAAGGCAATTAAAGCAAGATGCTAAAGCTATTGGAAG 1033
962 AGAGCTCAAGCGAGAGCAAGGCAATTAAAGCAAGATGCTAAAGCTATTGGAAG 1021
1034 CTTCACATGAGCACTGGAAGAAAGAGAGCAAGAAATTTGGCTGAGTGAATGTTGT 1093
1022 CTTCACATGAGCACTGGAAGAAAGAGAGCAAGAAATTTGGCTGAGTGAATGTTGT 1081
1094 GCTGTGCTCAATGAATTTGATTAACCAATGATTGTAAGCATTTAAAGTCAACAT 1153
1082 GCTGTGCTCAATGAATTTGATTAACCAATGATTGTAAGCATTTAAAGTCAACAT 1141
1154 ATTTTCATGAAGCACTGTTGAACCAATGCTGTTAAACCAAGCTGCCCATGTC 1213
1142 ATTTTCATGAAGCACTGTTGAACCAATGCTGTTAAACCAAGCTGCCCATGTC 1201
1214 AAATGTGACATCTCAACCTTTGGGAATTGAAGTGAATGTAAGTGAATGATGCT 1273
1202 AAATGTGACATCTCAACCTTTGGGAATTGAAGTGAATGTAAGTGAATGATGCT 1261
1274 TTAAGAGCTGCTGATTCACCAATTAATCTAATGAGCTCTCTCCATGAGAGTAAT 1333
1262 TTAAGAGCTGCTGATTCACCAATTAATCTAATGAGCTCTCTCCATGAGAGTAAT 1321
1334 CCGACGAGACCGGATCATCTGATATGCTTCACTGTAACAGAGACGATGAACGCTCTG 1393
1322 CCGACGAGACCGGATCATCTGATATGCTTCACTGTAACAGAGACGATGAACGCTCTG 1381
1394 GAGGAAACCTGAGTGAACAATTAAGAAAGCTCACTGCTGTAACCAATGAAGCAATTC 1453
1382 GAGGAAACCTGAGTGAACAATTAAGAAAGCTCACTGCTGTAACCAATGAAGCAATTC 1441
1454 GTGCAAGTGAATGTTATCTCATGATGTAACAACCACTTTGAAGAGAGCAACTCTCT 1513
1442 GTGCAAGTGAATGTTATCTCATGATGTAACAACCACTTTGAAGAGAGCAACTCTCT 1501
1514 AATCAAGAGCTGCTGTGGAAGAAATTAATCTTAAATCTGTGAATTTGAAGAACTGA 1573

Db 1502 AATCAAGAGACTGCTGTGGAGAAATTAAATCTGTAATATGAAAGAACTTGA 1561
 QY 1574 ACCATTAGTAATTAACAAGACTGCAATCAGGGCTTCTTATTAATTAATGGAATA 1633
 Db 1562 ACCATTAGTAATTAACAAGACTGCAATCAGGGCTTCTTATTAATTAATGGAATA 1621
 QY 1634 TTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1693
 Db 1622 TTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1681
 QY 1694 AATGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1753
 Db 1682 A--TGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1738
 QY 1754 CATAGGCAAGTTCTCTCGAG 1774
 Db 1739 CATAGGCAAGTTCTCTCGAG 1759
 RESULT 3
 AX576252
 LOCUS Sequence 33 from Patent WO0162922. 2876 bp DNA linear PAT 08-JAN-2003
 DEFINITION AX576252
 AX576252 GI:27646003
 VERSION AX576252.1 GI:27646003
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 Banville, S.C., Greenawald, L.B., Lincoln, S.E., Stockbrecher, T.K., Amshley, S., Chang, S.C., Chen, P., Dam, T.C., Liu, T.F., Rosen, B.H., Russo, F.D., D'Sa, S.A., Spiro, P.A., Bradley, D.L., Chen, A., Cohen, H.J., Daffo, A., Daniels, S.E., Dufour, G.E., Flores, V., Fong, W.T., Hodgson, D.M., Jackson, S., Jones, A.L., Panzer, S.R., Roseberry, A.M., Shah, P., Wright, R.J., Yap, P.B., Yu, J.Y., Bratcher, S.R., Chalup, M.S., Dahl, C.R., and Hillman, J.L., Polypeptides and corresponding molecules for disease detection and treatment Patent: WO 0162922-A 33 30-AUG-2001;
 JOURNAL Incyte Genomics, Inc. (US)
 FEATURES Location/Qualifiers
 source 1..2876 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /note="Incyte ID No: LI:757439.1:2000MAV01"
 BASE COUNT 840 a 566 c 642 g 825 t 3 others
 ORIGIN
 Query Match 80.7%; Score 1431.4; DB 6; Length 2876;
 Best local Similarity 96.5%; Pred. No. 0;
 Matches 1670; Conservative 0; Mismatches 28; Indels 33; Gaps 19;
 QY 74 CCAGAGAGCTGCATCTGCGGCAACCTGTGTGTAAGCTACGTCCTCTGAC-TCCGAC 132
 Db 33 CCAGAGAGCTGCATCTGCGGCAACCTGTGTGTAAGCTACGTCCTCTGAC 92
 QY 133 GTAGCTGCGAGCTCCCGAGCTGCACTCCATCTCCCTCCCACTGGGGCCGACCTGCTAA 132
 Db 93 GTAGCTGCGAGCTCCCGAGCTGCACTCCATCTCCCTCCCACTGGGGCCGACCTGCTAA 132
 QY 193 GACCAAGGTCCTGCGCAACGCTGAGAGGGGCGGTCCAGAGGGCGCTAAGGAACTCGAG 252
 Db 153 GACCAAGGTCCTGCGCAACGCTGAGAGGGGCGGTCCAGAGGGCGCTAAGGAACTCGAG 212
 QY 253 CGCGCCGCGCATGGGGCGCGGCTCGGGGCTCTCTGCGCGGGGCTGCGGGCTT 312
 Db 213 CGCGCCGCGCATGGGGCGCGGCTCGGGGCTCTCTGCGCGGGGCTGCGGGCTT 272
 QY 313 TTCAGATTGCTGAGTGTGCTCTGCTGCGCCCTGAGTCCGAGGCAACCGGTTCCCG 372

Db 273 TTCAGATTGCTGAGTGTGCTCTGCTGCGCCCTGAGTCCGAGGCAACCGGTTCCCG 332
 QY 373 GGGGGCTGAAGAGTGTGAAACCGCGTAACCTCAAGCTGTCTCGGGGCTTCCGACAACGG 432
 Db 333 GGGGGCTGAAGAGTGTGAAACCGCGTAACCTCAAGCTGTCTCGGGGCTTCCGACAACGG 392
 QY 433 AGTGAACGTAAGGTTGGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 492
 Db 393 AGTGAACGTAAGGTTGGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 452
 QY 493 GAGGCTGTGCTGAGGCTCTGTGACCGCCGACGAGGCGCGGAGCGCTTAAGCGCTGAA 552
 Db 453 GAGGCTGTGCTGAGGCTCTGTGACCGCCGACGAGGCGCGGAGCGCTTAAGCGCTGAA 512
 QY 553 CCGGCAACGAAATTCAGGAGTCCAGGCTTGGGGAGAGCAACGTCAGTCTGTGGT 612
 Db 513 CCGGCAACGAAATTCAGGAGTCCAGGCTTGGGGAGAGCAACGTCAGTCTGTGGT 572
 QY 613 GGCCCTCATCAACGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 672
 Db 573 GGCCCTCATCAACGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 632
 QY 673 GAAATGGGCTGTGAGACCGCTCATCTTTAACTCCCGGAGACCGGAAATGAGTATGAC 732
 Db 633 GAAATGGGCTGTGAGACCGCTCATCTTTAACTCCCGGAGACCGGAAATGAGTATGAC 692
 QY 733 CATGCTCAACCGGAGTGAAGACATTTGTCATCATGA-TGCGAATCTGAAGGCA 791
 Db 693 CATGCTCAACCGGAGTGAAGACATTTGTCATCATGA-TGCGAATCTGAAGGCA 752
 QY 792 CAAAAATTTGCAATCTTTTCAAGAGGAGTCAAGTGAAGAGTCAAGTGAAGAGTGAAGAG 851
 Db 753 CAAAAATTTGCAATCTTTTCAAGAGGAGTCAAGTGAAGAGTCAAGTGAAGAGTGAAGAG 812
 QY 852 AAAAAATGAGGCTTGGAGTCACTATTCATTTTTCG-TTCTGTGCTTTT 910
 Db 813 AAAAAATGAGGCTTGGAGTCACTATTCATTTTTCG-TTCTGTGCTTTT 872
 QY 911 ATTATTAAGGCTGAGGAGTCACTATTCATTTTTCG-TTCTGTGCTTTT 969
 Db 873 ATTATTAAGGCTGAGGAGTCACTATTCATTTTTCG-TTCTGTGCTTTT 932
 QY 970 TGCAAGAGCTCAACGAGAGAGAGG-CATTAAAGAGAGTCAATTAAGAGTCAATTAAG 1028
 Db 933 TGCAAGAGCTCAACGAGAGAGAGG-CATTAAAGAGAGTCAATTAAGAGTCAATTAAG 992
 QY 1029 GAGGC-TTCACTACGCACTGAAACAGAGGCA-GGAATTTGCGCTGAT-GGAG 1085
 Db 933 GAGGC-TTCACTACGCACTGAAACAGAGGCA-GGAATTTGCGCTGAT-GGAG 1052
 QY 1086 ATAGTTGCTGTGATGAAATTTGATTAACCAATGATTTGTAACGATCTTAAGT 1145
 Db 1053 ATAGTTGCTGTGATGAAATTTGATTAACCAATGATTTGTAACGATCTTAAGT 1112
 QY 1146 GCAACCAATTTTCCATTAAGCATGTGTGAACCATGCTGTTAACACAGACTTGC 1205
 Db 1113 GCAACCAATTTTCCATTAAGCATGTGTGAACCATGCTGTTAACACAGACTTGC 1172
 QY 1206 CCATGTGCAAAATGTAATACTCAAGCTTGGGAAATGAGTGAAGTGAAGTGAAT 1265
 Db 1173 CCATGTGCAAAATGTAATACTCAAGCTTGGGAAATGAGTGAAGTGAAGTGAAT 1232
 QY 1266 CAGTGT-CTTTCAAGTCCCTGT---ATCCAAATGAATCTAATAGTCTCTCCCA 1320
 Db 1233 CAGTGTCTTTCAAGTCTCTCTAATATCCAAATGAATCTAATAGTCTCTCTCCCA 1292
 QY 1331 TGAAGGAAATATGCAAGAGAGACCGGATCACTGGAATGCTCACTAAGGAGAGAG 1380
 Db 1293 TGAAGGAAATATGCAAGAGAGACCGGATCACTGGAATGCTCACTAAGGAGAGAG 1352
 QY 1381 TGAACCGCC-TCGAGAGAGACAGCTGAGTCAAC-AAATGAAGTCAAGCTGAT-AAA 1437
 Db 1353 TGAACCGCCCTTGAAGAGAGACAGCTGAGTCAAC-AAATGAAGTCAAGCTGAT-AAA 1412

QY 1438 CCATGAACCAATTCCT-GTGGAGAGATGTTATTCCTCATGTCAGCAACCAAC----- 1492
 Db 1413 CCATGAACCAATTCCTGTCGAGAGATGTTATTCCTCATGTCAGCAACCAACCTTN 1472
 QY 1493 ---TTTGAAGAGAGC---GAACTCCTCATGTCAGAGAGCTGTTGAGAAATTAATC 1545
 Db 1473 TTTTGGAGAGAGAGCTGGAAGAACTCATGTCAGAGAGCTGTTGAGAAATTAATC 1532
 QY 1546 TTAAGATCTGTTAATAGAGAAAC--TTGAGACCTATGTTAATAGAGAACTGTCAGAG 1603
 Db 1533 TTAAGATCTGTTAATAGAGAACTGTCAGAGAACTGTTAATAGAGAACTGTCAGAG 1592
 QY 1604 GGCCTAGTTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1663
 Db 1593 GGCCTAGTTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1652
 QY 1664 CTGAGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1723
 Db 1653 CTGAGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1710
 QY 1724 TTTTTCACCAAACTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1774
 Db 1711 TTTTTCACCAAACTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1760
 RESULT 4
 AF394689 1297 bp mRNA linear PRI 15-MAY-2003
 LOCUS Homo sapiens GRAIL (GRAIL) mRNA, complete cds.
 DEFINITION AF394689
 ACCESSION AF394689
 VERSION AF394689.1 GI:15011451
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS 1 (bases 1 to 1297)
 TITLE Anandaababathi, N., Ford, G.S., Bloom, D., Holmes, C., Paragas, V.,
 Sercoy, C., Skrenta, H., Hollemor, N., Pachman, C.G. and Soares, J.
 GRAIL: an E3 ubiquitin ligase that inhibits cytokine gene
 transcription is expressed in anergic CD4+ T cells
 JOURNAL Immunity 18 (4), 535-547 (2003)
 MEDLINE 12705856
 PUBMED 12705856
 REFERENCE
 AUTHORS 2 (bases 1 to 1297)
 TITLE Soares, J.R. and Pachman, G.
 GRAIL: a novel ring finger protein upregulated in anergic T cells
 JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 1297)
 AUTHORS Soares, J.R. and Pachman, G.
 TITLE Direct Submission
 JOURNAL Submitted (22-JUN-2001) Pathology, Stanford University, 800 Welch
 Rd, Palo Alto, CA 94304, USA
 FEATURES
 source
 1..1297
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="X"
 1..1297
 /gene="GRAIL"
 11..1297
 /note="E3, ring finger ubiquitin ligase; regulated in
 anergy-induced lymphocytes"
 /codon_start=1
 /product="GRAIL"
 /protein_id="AAK7554.1"
 /db_xref="GI:15011452"
 /translation="MGLPLGAGVSGRGGGSRLLMCFLLASPOAGSGAEAVT
 AVLVSRVPHGTGRTVWELSERGVGDSPLPAGVAVPPDGPALNACPHPTN
 TVPTVMSITVQVSWLALIQGGGCTPADKHLHAEKAGAVTFNPGTNEVTPMSH

PEADVAINIAGLKGKTLIOGIGQVTVIEVKGKHPVNHYSIFVSVFEFI
 TYAGGYFIFYSAXRLKXOSKOROLRADAKKXIGLOJRTKODKEIGPDS
 CAVCELYKNDLVRLTGNHIFHTCTGECGMCDCILKALGIEVDEG
 SVSLQVSNSEISNSASHERHNSRTRASSVAVSGDPLPLEHNVSTHESLQVNV
 HEANSVADVTIHYDPTFEDEPTNPTAVREIKS"
 BASE COUNT 324 a 311 c 355 g 303 t 4 others
 ORIGIN
 Query Match 72.2%; Score 1280.8; DB 9; Length 1297;
 Best Local Similarity 99.1%; Pred. No. 1e-285;
 Matches 1285; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
 QY 253 CGCGCGCGCATGAGGCGCGCGCTTGGGCGCGGCTTCTGCGCGGCTGCGGCTT 312
 Db 1 CGCGCGCGCATGAGGCGCGCGCTTGGGCGCGGCTTCTGCGCGGCTGCGGCTT 60
 QY 313 TTCCAGATGCTGAGCATGAGCTTCCGCTGAGCGCTGAGCGGAGGACCGGCTTCCG 372
 Db 61 TTCCAGATGCTGAGCATGAGCTTCCGCTGAGCGCTGAGCGGAGGACCGGCTTCCG 120
 QY 373 GGGAGCTGAGAGAGTGTGACCGCGTACCTGACCTGCTGCGGCTTCCGACAGG 432
 Db 121 GGGAGCTGAGAGAGTGTGACCGCGTACCTGACCTGCTGCGGCTTCCGACAGG 180
 QY 433 AGTGAACGCTAGGCTGTGAGAGCTGAGAGAGGCGCTGTACGAGACCTGCGCT 492
 Db 181 AGTGAACGCTAGGCTGTGAGAGCTGAGAGAGGCGCTGTACGAGACCTGCGCT 240
 QY 493 GGAAGCTGTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 552
 Db 241 GGAAGCTGTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 QY 553 CCGGCTGAGATTTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 612
 Db 301 CCGGCTGAGATTTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 QY 613 GGCCTCATTCACAGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 672
 Db 361 GGCCTCATTCACAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 QY 673 GAGATGAGGCTGTGAGCGCTCATCTTAACTTCCCGGAGCCGAGATGAGGCTATCC 732
 Db 421 GAGATGAGGCTGTGAGCGCTCATCTTAACTTCCCGGAGCCGAGATGAGGCTATCC 480
 QY 733 CATGCTCACCGGCTGTGAGAGCATGTGTGATCATGTGTGATCATGTGTGATCATGT 792
 Db 481 CATGCTCACCGGCTGTGAGAGCATGTGTGATCATGTGTGATCATGTGTGATCATGT 540
 QY 793 AAAAATTCGATCTATTAAGAGAGCATTAAGAGAGCATTAAGAGAGCATTAAGAGAG 852
 Db 541 AAAAATTCGATCTATTAAGAGAGCATTAAGAGAGCATTAAGAGAGCATTAAGAGAG 600
 QY 853 AAAAATTCGATCTATTAAGAGAGCATTAAGAGAGCATTAAGAGAGCATTAAGAGAG 912
 Db 601 AAAAATTCGATCTATTAAGAGAGCATTAAGAGAGCATTAAGAGAGCATTAAGAGAG 660
 QY 913 TATTAAGAGAGAGCATTAAGAGAGCATTAAGAGAGCATTAAGAGAGCATTAAGAGAG 972
 Db 661 TATTAAGAGAGAGCATTAAGAGAGCATTAAGAGAGCATTAAGAGAGCATTAAGAGAG 720
 QY 973 AAGAGCTCAAG 1032
 Db 721 AAGAGCTCAAG 780
 QY 1033 GCTTCAAG 1092
 Db 781 GCTTCAAG 840
 QY 1093 TCGTGTGATGATTAAG 1152
 Db 841 TCGTGTGATGATTAAG 900
 QY 1153 TATTTTCATTAAG 1212

```

Db      901 TATTTCCATAGACAGTGTGTACCCATGCTGTGTAGAACACAGACCTTGCCTCCCATGTG 960
QY      1213 CAATATGACATCTCAACAGCTTTGGCAATTTAGGTGATGTTGAAGATGATCAGTCTC 1272
Db      961 CAATGTGTGACATCAACAGCTTTGGCAATTTAGGTGATGTTGAAGATGATCAGTCTC 1020
QY      1273 TTACAAAGTCCCTGTATTCGAATTAATCTAATGTAGTCCCTCTCCCATGAGAGATTA 1332
Db      1021 TTTACAGTCCCTGTATTCGAATTAATCTAATGTAGTCCCTCTCCCATGAGAGATTA 1080
QY      1333 TCGCAGCGAGACCGCATCATCTGATATTCCTTCACTACAGGGAACAGATGAACCGCTCT 1392
Db      1081 TCGCAGCGAGACCGCATCATCTGATATTCCTTCACTACAGGGAACAGATGAACCGCTCT 1140
QY      1393 GAGGAGACCGCTGACGATCAACAATTAAGTCTACAGTGTGTAACCAAGAGCAATTC 1452
Db      1141 GAGGAGACCGCTGACGATCAACAATTAAGTCTACAGTGTGTAACCAAGAGCAATTC 1200
QY      1453 TGTGCAAGTGAATGTATTCCTCATCTGATGACACCAACCTTTGAAGAGAGCAATCTCC 1512
Db      1201 TGTGCAAGTGAATGTATTCCTCATCTGATGACACCAACCTTTGAAGAGAGCAATCTCC 1260
QY      1513 TAATCAAGAGACTGCTGTGTGAGAAATTAATCTTAA 1549
Db      1261 TAATCAAGAGACTGCTGTGTGAGAAATTAATCTTAA 1297

RESULT 5
LOCUS   BC010477 1928 bp mRNA linear ROD 10-JUN-2003
DEFINITION Mus musculus ring finger protein 128, mRNA (CDNA clone MGC:19395
ACCESSION BC010477
VERSION   BC010477.1 GI:14714672
KEYWORDS  MGC.
SOURCE    Mus musculus (house mouse)
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,K.H., Scheaffer,C.F., Bhat,N.K.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Schaefer,T.E., Brownstein,M.U., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullaly,S.J., Bosak,S.A., McMan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Woley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Faney,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shchepochko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalins,D.B.,
Schnerch,A., Schein,J.B., Jones,S.J. and Marra,K.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

TITLE
JOURNAL   MEDLINE
PUBMED    22388257
12477932
AUTHORS    Strausberg,R.
TITLE      Direct Submission
JOURNAL    Submitted (10-JUN-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK     MGC Project URL: http://mgc.nci.nih.gov
COMMENT    Contact: MGC help desk

```

```

FEATURES
source
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129, C57BL/6J, FVB/N"
/db_xref="taxon:10090"
/clone="MGC:19395 IMAGE:3154433"
/tissue_type="mammary tumor. Brca1-/-fl; MMTV-Cre model. 10
months old, gross tissue."
/clone_lib="NCI CGAP Mam3"
/lab_host="DH10B"
/note="Vector: pCMV-Sport6"
1. 1928
/gene="Rnf128"
/note="synonyms: 1300002C13Rlk, GRAL1"
/db_xref="LocusID:66889"
/db_xref="MGI:1914139"
176. 1462
/codon_start=1
/product="Riken cDNA 1300002C13"
/protein_id="P4M10477.1"
/db_xref="GI:14714673"
/locus="LocusID:66889"
/translaton="MGPPIGVYCGGCGGAGATLAWCFLLALSPHAPGSRGAERAVT
AYLVNRYERGVARTMSEGVGDSLEPVSGLVDPGQALNANCPNTF
TPPTWVGSTVOVMALILORGCGTADLHILASRSGAGLVNPGRNIVPMH
PKGQDVALTMGNKGTILOSTORIGVMTIVGKKRPMVNVHVSIFPVSPFII
TAATVGFPIFYSARILNARARQRKROKAPAKALOKLRTIVQSPKEIGPQDS
CAVCLILEVNMVRLNCHIFHKTCPVPLHRTQPMWCKDILKALGIEVDYDG
SVSLQVPEVNEASVNSPHEBDSRSTASGVASVQADDEPLLEHVSANMLDLVN
HEANSVAVDVVPHVNDPFEEDETPDORAAVREIKS"
BASE COUNT 482 a 466 c 498 g 482 t
ORIGIN
Query Match 67.9%; Score 1205.2; DB 10; Length 1928;
Best Local Similarity 89.6%; Pred. No. 3,7e-268;
Matches 1364; Conservative 0; Mismatches 149; Indels 10; Gaps 6;
QY 252 GAGCGCGGCGCATGAGGCGCGCGCTGAGGCGCGGAGTCTCTGCGCGGAGTGGCGGCT 311
Db 165 GAGTGTGCGGCGCATGAGGCGCGCGCGCTGAGGCGCGGAGTCTCTGCGCGGAGTGGCGG 224
QY 312 TTTCAGATTGCTGGCATGTGTCTTCTGTGCGGCTGAGTCCGAGGACCGGTTCCG 371
Db 225 CTGCGCGGCGCATGAGGCGCGCGCTGAGTCTCTGAGTCCGAGGACCGGCGGTTCCG 284
QY 372 GAGGAGCGGAGAGAGTGAAGCGGATCCGATCAACGATGCTGAGGAGTTCCGACACG 431
Db 285 GAGGAGCGGAGAGCGGTGTGAACCGGATCCGATCAACGATGCTGAGGAGTTCCGACACG 344
QY 432 GAGTGAACCGGAGCGGTGTGAACCGGATCCGATCAACGATGCTGAGGAGTTCCGACACG 491
Db 345 GAGTGAACCGGAGCGGTGTGAACCGGATCCGATCAACGATGCTGAGGAGTTCCGACACG 404
QY 492 TGAAGCGTGTGAGGCTGTGTGTATCCGCGGACCGGCGCGGCGCTTAAACCGCTGT 551

```

Email: cgabs-remail.nih.gov
Tissue Procurement: Lothar Henniphausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gumaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsged, H.,
Kovis, C.R., Sneed, A.O., Martin, R.G., Muzny, D.M., Nambavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: IRK Plate: 23 Row: b Column: 12
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 12563610.
Location/Qualifiers

BASE COUNT	477 a	466 c	506 g	487 c
ORIGIN				

DY 252 GGCAGCCGCGCATGGAGGCCGCGCCTGGAGACCGGGGTTCCTGCACCGGGTGGACTG 311
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 176 GAATGCGCGCAATTGGAGCCGCGCCGCGGAATCGGAGGTAACTACGCCGCGCGAGCTCGGAG 235
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

22	CGGCGGCGGCTGTAAGCACTGTGGAAACCGCGTAACCTCAACGTCCTCGGCGGGGTTCCGCAACAAG	431
Db	236 CTGCGCCGCTACTGSGCTTGAGCTTCTCTTCTTGCTCTGAGTCCGCAACGCGCCCGGATTC	295
Qy	372 GGGGCGGCTGTAAGCACTGTGGAAACCGCGTAACCTCAACGTCCTCGGCGGGGTTCCGCAACAAG	431
Db	296 GGGGAGCGGAAGCCGCTGTGAGAACCGCGTAACCTCAACGTCCTCGGCGGGGTTCCGCAACAAG	355
Qy	432 GAGTGAACCGTACCGGTGTGGAGCTGAGCGAAGGAGGGGAGTGAACGCGACAGGACTCGACCG	491
Db	356 GAGTGAACCGTACCGGTGTGGAGCTGAGCGAAGGAGGGGAGTGAACGCGACAGGACTCGACCG	415
Qy	492 TGGAGCCCTGTGGAGCTGTGGAGCTCTGTGTAACCGCCGACCGGAGCCGGGGGCGCTTAACGCGCTGA	551
Db	416 TGGAGCCCGCTCTCGGGGCTCTGTGTAACCGCCGACCGGAGCCGGGGGCGCTTAACGCGCTGA	475
Qy	552 ACCCGCAACAATTTTCAAGGTGTGCCACAGGTTTGGGGAAAGAACCTGTGAAATCTCTTGGT	611
Db	476 ACCCGCAACAATTTTCAAGGTGTGCCACAGGTTTGGGGAAAGAACCTGTGAAATCTCTTGGT	535
Qy	612 TGGCCCTCATATCAACCGCGCGGGGCTGCACTCTCGACAGACAATCAATCATCTGCGCTATG	671
Db	536 TGGCCCTCATATCAACCGCGGTGGAGGCTGCACTCTCGCGAGAGATCATCTGCGGCTGAG	595
Qy	672 AGAATGTGGCGGTCTGAGACCGCTCATCTTTAACTTCCCCGGGAAACCCGAAAGAGATCAAC	731
Db	596 AGAAGAGGGGCTTTGGAGGGGCTCATCTTTAACTTCTCGGGAACCCGAAAGAGATCAAC	655
Qy	732 CCATGCTCAACCGGGTGCAGTGAACATTTGTCATCAATGATCGGCAATCTGAAAGCA	791
Db	656 CCATGCTCAACCGGGTGTCTGGGACATTTGTCATCAATGATTTGGCAATCTGAAAGGA	715
Qy	792 CAAAATTTGCAATCTATTCAAGAGGCAACAATGCAATGCTCATAGAAGTAGGA	851
Db	716 CAAAATTTGCAAGTCTATTCAAAAGGCAATCAATGCAATGCTCATAGAAGTAGGA	775
Qy	852 AAAACATGAGCCCTTGGGAGTCACTATTCATATTTTTCGTTCTGTGTCCTTTTGA	911
Db	776 AAAACATGAGCCCTTGGGAGTCACTATTCATATTTTCTGTGTCCTTTTGA	835
Qy	912 TTATTACGGCGGCAACTGTGGGCTTTTATCTTTATCTTGTCTGAAGCTACGGAATG	971
Db	836 TTATTACGGCGCAACCGTGGGCTTTTATCTTGTCTGTGAAAGTTACGAATAG	895
Qy	972 CAAAGACTCAAGAGCGGAAGCAGAGGCACTTAAAGCAGATCTTAAAGTAGTATGGA	1031
Db	896 CAAAGACTCAAGAGCGGAAGCAGAGGCACTTAAAGCAGATCTTAAAGTAGTATGGA	955
Qy	1032 GGGCTTCAACTACGCACTGAAAACAAGAGCAAGAAAATTGGCCCTGATGAGATGTT	1091
Db	956 AAGCTTCAAGCTCTCGACCTTGAACAAGAGCAAGAAAATTGGCCCTGATGAGATGAT	1015
Qy	1092 GAGCTGTGTGATGATGATGATGATTAACAATATGTTTGGTACGATCTTAACGTGCAAC	1151
Db	1016 GTGCTGTGTGATGATGATGATCTTAAGCCAAATATTTGGTGGCATCTTAACGTGCAATC	1075
Qy	1152 AATATTTTCATAGACATGTGTGACCAATGGCTGTAAACAACAATGTCGCCATGT	1211

Db	1076	ATATTTTCCATTAAGACATGCTGACCCCGGCTTTTGAAACACAGSACTTGCCCATGT	1135
Qy	1212	GCAAAATGACATCTCAAGCTTTGGGAATGAGGTGATGTGGAAGATGATCATGT	1271
Db	1136	GCAAGTGTGCAATCTCCAAAGCTCTGGGAATTGAAGTGAATGTTGAAGTGCAGTGT	1199
Qy	1272	CTTAAACAAGCCCTGATTCGAAATGAATATCTAATAGAGCCCTCCCATGGAAGAGATA	1331
Db	1136	CTTTACAGTCTCTGTTCTTAATAGCATCTAATCTGCTCTCCGCCAAGGAAGSACA	1255
Qy	1332	ATGCAACGAGAACCGCATCATCTGATGATGCTTCAGATAACAGGAAACGATGAACGCTCTC	1391
Db	1256	GTCACATGAGATGTGATCATCTGGAATATCTTCAGTACAGAGACGAGATGACCACTCTC	1315
Qy	1332	TGAGAGAACACGTGCAGTCAACAAATGGAAGCTACAGCTGTAAACATGAAAGCAAT	1451
Db	1316	TGGAGGAACATGGGCACTGACGATGCAATTAATTCAGCTGTAAACCATGAAAGCAAT	1375
Qy	1452	CTGAGCAGTGAATGTTATCTCATGTGACACCCAACTTTGAAGAAAGCAAACTC	1511
Db	1376	CTGAGCGGTGAGATGTTGTTCCCATATTTGACAAACCAACTTTGAAGAAAGCAAACTC	1435
Qy	1512	CTAATCAAGAGCTGTGTCGAGAAATTTAACTTT-AAAATCTGTGTAATTTGAAAATCT	1570
Db	1436	CTGATCAAGAGGAGGTGTGTGGAGATTTAACTTTAAATCTGTGTCAATTTGAAAATCT	1495
Qy	1571	TGAACCATTTAG-TAATAACAAGACTGCCAATCAAGGCTCAATTT-CTATTATTAATTTGG	1628
Db	1496	TGAACCTTTAGTTAACAACAAGACTGCCAATCAAGGCTCAATTTACTATTAATTAAGACTGG	1555
Qy	1628	ATAAATTTAATAAATAAGAGTGAATCTGAATAGTGTCAAGATGCTAATTTTGTGTATA	1688
Db	1556	GTAATA--GATAAACAAGATGATCTGAATGTGTGAGTATCTTAATTTTACTATA	1613
Qy	1688	GTTAAATGSGCTAAATAAATTTAACTGTAACTTTTTCACCAAACTCAATATATA	1748
Db	1613	GTTAAA--TGGTTAACAATTTAACCCAGATCCGTTTT--CCAACAATCTACATATAAG	1668
Qy	1748	TTTTTCAATAGGCAAGTTCTCT 1771	
Db	1669	TTTTTCAATAGGCAAGTTCTCT 1691	

```

RESULT 7
LOCUS      AY112656
DEFINITION Mus musculus E3 ubiquitin ligase (greul1) mRNA, complete cds.
VERSION    AY112656
KEYWORDS   GI:23477774
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
REFERENCE  1
AUTHORS    Fukuyota, M.; Metzger, J.; Chodura, C.; Craniata, V.; Vertebrata, E.; Euteleostomi, M.; Mammalia, E.; Eutheria, R.; Rodentia, S.; Sciuognathi, M.; Muridae, M.; Murinae, M.; Mus. 1 (bases 1 to 1513)
           Borchers, A.G.M., Hufcon, A.L., Eldridge, A.G., Jackson, P.K., Harland, R.M. and Baker, J.C.
           The E3 ubiquitin ligase GREUL1 Anteriorizes Ectoderm during Xenopus Development
           Dev. Biol. 251 (2), 395-408 (2002)
JOURNAL    22322926
MEDLINE    12453366
PubMed     12453366
REFERENCE  2 (bases 1 to 1513)
AUTHORS    Baker, J.C.
TITLE      Direct Submission
JOURNAL    Submitted (15-May-2002) Genetics, Stanford, 300 Pasteur Dr.,
           MS5120, Rm. M337, Stanford, CA 94305, USA
           Location/Qualifiers
           1..1513
           /organism="Mus musculus"
           /mol_type="mRNA"
           /strain="CD-1"
FEATURES
source

```


CDS

/mol_type="mRNA"
/strain="DBA/2"
/db_xref="taxon:10090"
/clone="1.3.7"
/cell_type="nerve T cell"
1.1287

/codon_start=1
/product="zinc ring finger-containing protein GRIL"
/protein_id="P43451.1"
/db_xref="gi:17046406"

/translation="MGPFGIGVCRGGCGCAALLMWTILASPHAPSGRKAHMT
AYLWAKVPEHETWNTWLSERVTGDSPLPVSGLVPPDGPALNACHPENP
TPTWSTVQVSMALLQGGGCTPADKTLHLSERAGAVINPFGTNEIPSH
PGADIVAMIGNIGKTKLIGSIOGIVMTLVGKSGVMHYSIFVSVSPFI
TAATVGFIFYSARLRNARAKSRKQKADAKKALIGLQRLTKGQDEIGPDS
CAVCIKYEKNDIVRLITCNHIFKTCVDPMLHEHTGCMCKDITLALIEVDYD
SVSLQVPSNENSTASPHEDSRSETPASGVAASGADPEPPEHROSANENLQVAV
HEANSVADVPHVDNPTPEDETPDEAAVREIKS"

BASE COUNT 308 a 320 c 359 g 300 t
ORIGIN

62.4%; Score 1106.8; DB 10; Length 1287;

Query Match Best Local Similarity 91.2%; Pred. No. 2.1e-245;
Matches 1174; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

263 ATGGGAGCCGCGCTGGGGCCGGGGTCTCTCCGCGCGTGGCGCTTTCCAGATTG 322
1 ATGGGAGCCGCGCGCGCGGATCGGGGTCTACTGCGCGCGCGCTGCGAGCTGCCGGCTA 60
323 CTGGCATGGTCTCTCTGCTGAGTCCGAGGCAACCGGTTCCCGGGGGCTGA 382
61 CTGGCTGGTCTCTCTCTGCTGAGTCCGAGGCAACCGGTTCCCGGGGGCTGA 120
383 GCGAGTGGACCGCGTACTCTAACTGCTGCGGGGTTCGCGACGCGGAGTAACTT 442
121 GCGGTGGACCGCGGTCTCTAACTGCTGCGGGGTTCGCGACGCGGAGTAACTT 180
443 AGCGTGGAGAGCTGAGCGAGGCGGCTGACGCGGAGTCTCGCGTGGAGCTGTG 502
181 AGCGTGGAGAGCTGAGCGAGGCGGCTGACGCGGAGTCTCGCGTGGAGCTGTG 240
503 GCTGGAGTCTGAGTACCGCGCGCGCGCGGCTTAACTGTAACCGCTGAACCGCAG 562
241 TCCGGGGTCTGAGTACCGCGCGCGCGCGGCTTAACTGTAACCGCTGAACCGCAG 300
563 AATTTCAGCGTCCCGCGGTTGGGAGAGCGCGGAGTCTGTTGGTGGCTCTATC 622
301 AATTTCAGCGTCCCGCGGTTGGGAGAGCGCGGAGTCTGTTGGTGGCTCTATC 360
623 CACGCGCGCGGCGTCACTGCGAGCAAGATCTGCTTATGAGAGTGGCG 682
361 CACGCGCGGCGTCACTGCGAGCAAGATCTGCTTATGAGAGTGGCG 420
683 TCTGAGCGCTGATCTTAACTTCCCGGAGCCGCAATGAGTCACTCCCATGCTTAC 742
421 TCTGAGCGCTGATCTTAACTTCCCGGAGCCGCAATGAGTCACTCCCATGCTTAC 480
743 CCGGTCAGTGAACATGTTGCAATGCGCAATCTGAAAGCAAAATTCG 802
481 CCGGTCAGTGAACATGTTGCAATGCGCAATCTGAAAGCAAAATTCG 540
803 CAATCTATCAAGAGCAATGAGTGAATGCTGAAAGTGGGAAAAATGTC 862
541 CAATCTATCAAGAGCAATGAGTGAATGCTGAAAGTGGGAAAAATGTC 600
863 CTTGGGAGTGAATCAATGATTTTTCGTTCTGCTGCTTTTATTAAGCGCG 922
601 CTTGGGAGTGAATCAATGATTTTTCGTTCTGCTGCTTTTATTAAGCGCG 660
923 GCACTGAGGCTATTTATCTTTATCTGCTGAGAGGCTAGCGAGTCAAGCTCAA 982
662 GCACTGAGGCTATTTATCTTTATCTGCTGAGAGTCAAGTCAAGAGCTCAA 720

QY 983 AGCAGGAGACAGAGGCAATTAAAGCAGTGTAAATAAGCTATGAGAGCTTCACTA 1042
DB 721 AGCAGGAGACAGAGGCAATTAAAGCAGTGTAAATAAGCTATGAGAGCTTCACTA 780
QY 1043 CGCAGCTGAAACAGAGAGCAAGAAATGAGCCCTGATGAGATGTTGCTGTGTC 1102
DB 781 CGCAGCTGAAACAGAGAGCAAGAAATGAGCCCTGATGAGATGTTGCTGTGTC 840
QY 1103 ATTGAATGTAATAACCAATGATTTGGAGCATCTTAAAGTCAACCAATTTTCAT 1162
DB 841 ATTGAATGTAATAACCAATGATTTGGAGCATCTTAAAGTCAACCAATTTTCAT 900
QY 1163 AAGACATGTGTGTAACCATGCTGTTAAACAGAGCTTCCCATGCAATGTAGC 1222
DB 901 AAGACATGTGTGTAACCATGCTTAAACAGAGCTTCCCATGCAATGTAGC 960
QY 1223 ATCTGTAAGCTTGGAGATTTGAGTGTGATGTTGAATGATGATGCTTAACTC 1282
DB 961 ATCTGTAAGCTTGGAGATTTGAGTGTGATGTTGAATGATGATGCTTAACTC 1020
QY 1283 CTTGATCCATGAAATATCTAAATAGTGTCTCTCCATGAGAGATATGCGACGAG 1342
DB 1021 CTTGATCCATGAAATATCTAAATAGTGTGTCTCTCCATGAGAGATATGCGACGAG 1080
QY 1343 ACCGATCATCTGATATCTTCAATGAGAGAGCAAGTACCGCTCTGAGAGAAC 1402
DB 1081 ACCGATCATCTGATATCTTCAATGAGAGAGCAAGTACCGCTCTGAGAGAAC 1140
QY 1403 GTGCACTGCAACAAATGAAATGCTACAGCTGTGTAACATGAGCAATTTCTGCACTG 1462
DB 1141 GTGCACTGCAACAAATGAAATGCTACAGCTGTGTAACATGAGCAATTTCTGCACTG 1200
QY 1463 GATGTTATCTCTCATGAGAGCAACCACTTTGAGAGAGCAATCTTAAATGAG 1522
DB 1201 GATGTTATCTCTCATGAGAGCAACCACTTTGAGAGAGCAATCTTAAATGAG 1260
QY 1523 ACTGCTGTGAGAAATTAATCTTAA 1549
DB 1261 ACTGCTGTGAGAAATTAATCTTAA 1287

RESULT 9

AK027169
LOCUS Homo sapiens cDNA: FLJ23516 fls, clone IMAGE4848.
DEFINITION AK027169
ACCESSION AK027169
VERSION 1
KEYWORDS oligo capping; fls (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE

1 (sites)
Kawakami, T., Noguchi, S., Itoh, T., Shigeta, K., Senba, T.,
Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Tanigami, A.,
Fujisawa, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hirao, M.,
Omori, Y., Ota, T., Suzuki, Y., Obayashi, M., Nishi, T., Shibahara, T.,
Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.
NEDO human cDNA sequencing project

TITLE

Unpublished

JOURNAL

2 (bases 1 to 1871)

REFERENCE

Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T., Isogai, T.,
Shibahara, T., Tanaka, T. and Nakamura, Y.
Direct Submission

AUTHORS

Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure Analysis, Human
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
Japan (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)

TITLE

NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing; Department of

COMMENT

Unpublished

COMMENT

Unpublished

Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).

FEATURES
source
1..1871
location/Qualifiers

CDS

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="LNG04848"
/issue_type="human lung"
/clone_id="LNG"
/note="Cloning vector pME18SFL3"
76..546
/note="unnamed protein product"
/codon_start=1
/protein_id="BA15682.1"
/db_xref="GI:10440233"
/translation="MEIVYLCAINCINQMTWISILNQHIFHKTCVDPWLEHRTCPM
CKQDIKALGIYDVDSASIQYPSNHSNSASHSRSTASGYSYQGTDS
PPLREHYQSTNBSLOIVNHSANSVAVDVIPHDNPTFSDPTFQGTVAEIKS"

BASE COUNT 670 a 277 c 321 g 603 t
ORIGIN

Query Match 40.3%; Score 714.4; DB 9; Length 1871;
Best Local Similarity 97.9%; Pred. No. 1.7e-154;
Matches 755; Conservative 0; Mismatches 12; Indels 4; Gaps 3;

1005 AGCGAGATGCTAAAGGCTATTGGAAGGCTTCAACTACGCACTGAAACAGAGACA 1064
1 AGCGAGATGCTAAAGGCTATTGGAAGGCTTCAACTACGCACTGAAACAGAGACA 60
1065 AGGAAATGGCCGATGAGAGAGATGCTGCTGATGATGATGATGATGATGATGATG 1124
61 AGGAAATGGCCGATGAGAGAGATGCTGCTGATGATGATGATGATGATGATGATGATG 120
1125 ATTGGTAC-GCATCTTACGCTGCAACCAATTTCCATTAAGACATGCTGACCATG 1183
121 ATTGGTACGACATCTTACGCTGCAACCAATTTCCATTAAGACATGCTGACCATG 180
1184 CTGTTAAACACAGACTGCCCCCATGTCACAAATGTCATCTCAAGCTTTGGGAATT 1243
181 CTGTTAAACACAGAGACTGCCCCCATGTCACAAATGTCATCTCAAGCTTTGGGAATT 240
1244 GAGTGTGATGTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1303
241 GAGTGTGATGTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
1304 AATAGTCTCTCTCCCATGGAAGATGATGATGATGATGATGATGATGATGATGATG 1363
301 AATAGTCTCTCTCCCATGGAAGATGATGATGATGATGATGATGATGATGATGATG 360
1364 TCAATGACAGAGACAGATGAGACGCTCCGAGAGACAGTGCATGCAACAAATGAAAGT 1423
361 TCAATGACAGAGACAGATGAGACGCTCCGAGAGACAGTGCATGCAACAAATGAAAGT 420
1424 CTACAGCTGTGAACATGAGACAAATTCGTGACAGTGTGATGATGATGATGATG 1483
421 CTACAGCTGTGAACATGAGACAAATTCGTGACAGTGTGATGATGATGATGATG 480
1484 AACCAACCTTTGAAGAGACGAAATCTCTATACAGACGCTGCTGCGAATTTAA 1543
481 AACCAACCTTTGAAGAGACGAAATCTCTATACAGACGCTGCTGCGAATTTAA 540
1544 TCTTAAATCTGTGTAATGAGAACTTGAACCATAGTATGATGATGATGATGATG 1603
541 TCTTAAATCTGTGTAATGAGAACTTGAACCATAGTATGATGATGATGATGATG 600
1604 GGCCTAGTTCTATTAATTAATGATGATGATGATGATGATGATGATGATGATG 1663
601 GGCCTAGTTCTATTAATTAATGATGATGATGATGATGATGATGATGATGATG 660
1664 CTACAGTACTAATATATGCTATGATGATGATGATGATGATGATGATGATGATG 1723

Db 661 CTCAGATGACATTAATTAATGATGATGATGATGATGATGATGATGATGATGATG 718
QY 1724 TTTTCCACCAAACTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1774
Db 719 TTTTCCCA CAATCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 768

RESULT 10
AL391315
LOCUS
DEFINITION Human DNA sequence from clone RP11-150F24 on chromosome X, complete
sequence.
ACCESSION AL391315
VERSION AL391315.22 GI:14575271
KEYWORDS HTG.
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 169812)

REFERENCE
AUTHORS Brown, A.
TITLE Direct Submission
JOURNAL Submitted (27-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
request: clonerequest@sanger.ac.uk
On Jun 28, 2001 this sequence version replaced gi:14455902.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Bm, EMBL; Sw,
SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
This sequence was generated from part of bacterial clone contigs of human
chromosome X, constructed by the Sanger Centre Chromosome X Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/ChrX
RP11-150F24 is from the library RPCI-11.1 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBAC3.6

This sequence is the entire insert of clone RP11-150F24. The true
left end of clone RP11-321G1 is at 68861 in this sequence. The true
right end of clone RP11-697G3 is at 161410 in this sequence.

FEATURES

source

1..169812
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="X"
/clone="RP11-150F24"
/clone_id="RPCI-11.1"

misc_feature

12..569
/note="match: GSS: Bm: A0375615"

repeat_region

94..165
/note="72 copies 36 mer 90% conserved"

repeat_region

512..729
/note="72R repeat: matches 3..238 of consensus"

repeat_region

1338..1638
/note="AluY repeat: matches 1..304 of consensus"
3599..3642
/note="12 repeat: matches 2707..2750 of consensus"

repeat_region 3722..3899
 /note="MIR repeat: matches 12. .171 of consensus"
 repeat_region 3915..4028
 /note="MIR repeat: matches 131. .238 of consensus"
 repeat_region 5121..5197
 /note="L2 repeat: matches 2611. .2710 of consensus"
 repeat_region 6734..7029
 /note="AluSx repeat: matches 1. .297 of consensus"
 repeat_region 7660..7955
 /note="AluSx repeat: matches 1. .308 of consensus"
 repeat_region 8891..8940
 /note="L1PA9 repeat: matches 6114. .6163 of consensus"
 repeat_region 10067..10370
 /note="AluSx repeat: matches 1. .297 of consensus"
 repeat_region 10411..10695
 /note="AluSx repeat: matches 1. .288 of consensus"
 repeat_region 10737..10879
 /note="MER3 repeat: matches 5. .154 of consensus"
 repeat_region 11078..11380
 /note="AluY repeat: matches 1. .300 of consensus"
 repeat_region 11571..12560
 /note="L1MD2 repeat: matches 5328. .6331 of consensus"
 repeat_region 13476..13794
 /note="L1MD1 repeat: matches 5837. .6165 of consensus"
 repeat_region 13906..14041
 /note="L2 repeat: matches 2536. .2671 of consensus"
 repeat_region 14106..14411
 /note="AluY repeat: matches 2. .306 of consensus"
 repeat_region 16413..16704
 /note="AluSx repeat: matches 1. .295 of consensus"
 repeat_region 17566..17697
 /note="FLAM C repeat: matches 1. .132 of consensus"
 repeat_region 17839..18080
 /note="MIR repeat: matches 3. .260 of consensus"
 repeat_region 18961..19153
 /note="MIR repeat: matches 30. .262 of consensus"
 repeat_region 19562..19658
 /note="MER3 repeat: matches 12. .109 of consensus"
 repeat_region 20326..20413
 /note="L1MD4 repeat: matches 7861. .7948 of consensus"
 repeat_region 20471..20754
 /note="AluSx repeat: matches 1. .286 of consensus"
 repeat_region 20755..20782
 /note="L14 copies 2 mer ta 8% conserved"
 repeat_region 20939..21184
 /note="L2 repeat: matches 2529. .2729 of consensus"
 repeat_region 21194..21952
 /note="L2 repeat: matches 1730. .2546 of consensus"
 repeat_region 22817..23035
 /note="MIR repeat: matches 15. .238 of consensus"
 repeat_region 23641..24552
 /note="CpG island"
 /evidence="not_experimental"
 repeat_region 24009..24384
 /note="L1MD repeat: matches 310. .694 of consensus"
 repeat_region 25462..25711
 /note="match: GSS: Em:AQ923340"
 repeat_region 25838..25939
 /note="MIR repeat: matches 2. .256 of consensus"
 repeat_region 26420..26566
 /note="MIR repeat: matches 6. .106 of consensus"
 repeat_region 28132..28302
 /note="MIR repeat: matches 2. .151 of consensus"
 repeat_region 29706..29835
 /note="L2 repeat: matches 64. .238 of consensus"
 repeat_region 30126..30458
 /note="MER2 repeat: matches 2. .344 of consensus"
 repeat_region 30873..31184
 /note="AluSx repeat: matches 1. .312 of consensus"
 repeat_region 34418..34503
 /note="L1PA3 repeat: matches 6061. .6146 of consensus"

repeat_region 34518..34851
 /note="MER2 repeat: matches 1. .344 of consensus"
 repeat_region 35308..35457
 /note="MIR repeat: matches 1. .162 of consensus"
 repeat_region 35828..36039
 /note="MIR repeat: matches 7. .226 of consensus"
 repeat_region 36274..36349
 /note="L19 copies 4 mer ttcg 76% conserved"
 repeat_region 36353..36642
 /note="AluSx repeat: matches 1. .291 of consensus"
 repeat_region 37753..37919
 /note="FRAM repeat: matches -1. .165 of consensus"
 repeat_region 38928..39134
 /note="L2 repeat: matches 2467. .2703 of consensus"
 repeat_region 39127..39319
 /note="L2 repeat: matches 2031. .2228 of consensus"
 repeat_region 39346..39726
 /note="MER3 repeat: matches 1. .186 of consensus"
 repeat_region 39738..40910
 /note="Charliela repeat: matches 1. .1258 of consensus"
 repeat_region 40911..41435
 /note="L1PA5 repeat: matches 5616. .6143 of consensus"
 repeat_region 41436..41674
 /note="Charliela repeat: matches 1258. .207 of consensus"
 repeat_region 41675..41942
 /note="AluY repeat: matches 1. .283 of consensus"
 repeat_region 41943..42098
 /note="Charliela repeat: matches 207. .1455 of consensus"
 repeat_region 42634..42699
 /note="WSRP repeat: matches 356. .426 of consensus"
 repeat_region 43609..43659
 /note="L2 repeat: matches 2696. .2746 of consensus"
 repeat_region 45090..45187
 /note="L2 repeat: matches 2606. .2705 of consensus"
 repeat_region 45672..46211
 /note="L1MD4 repeat: matches 5700. .6295 of consensus"
 repeat_region 46808..47107
 /note="AluY repeat: matches 1. .301 of consensus"
 repeat_region 47190..47271
 /note="HY1 repeat: matches 28. .111 of consensus"
 repeat_region 47355..47455
 /note="MER3 repeat: matches 5. .734 of consensus"
 repeat_region 47525..47693
 /note="MIR repeat: matches 3. .196 of consensus"
 repeat_region 47826..47978
 /note="match: GSS: Em:AQ136922"
 repeat_region 47985..48294
 /note="match: GSS: Em:AQ136922"
 repeat_region 48203..50340
 /note="L1M1 repeat: matches -1389. .1172 of consensus"
 repeat_region 50335..51398
 /note="L1PA2 repeat: matches 1679. .2716 of consensus"
 repeat_region 51399..51743
 /note="AluY repeat: matches 1. .305 of consensus"
 repeat_region 51744..52940
 /note="L1PA2 repeat: matches 2716. .3982 of consensus"
 repeat_region 52917..54922
 /note="L1PA2 repeat: matches 4138. .6146 of consensus"
 repeat_region 54936..57273
 /note="L1MD3 repeat: matches 3351. .6304 of consensus"
 repeat_region 57453..58059
 /note="L1MD3 repeat: matches 5276. .5875 of consensus"
 repeat_region 58056..58804

Query Match 38.0% Score 674.4 DB 9: Length 169812;
 Best Local Similarity 99.9% Pred.No.9.1e-145 Indels 0 Gaps 0;
 Matches 675; Conservative 0; Mismatches 1;

QY 74 CGAGAGCTGATCTGGGCACTGTGTGCTGAGCGCTACGATCTCTCTGCTCGAGC 133
 Db 23909 CGAGAGCTGATCTGGGCACTGTGTGCTGAGCGCTACGATCTCTCTGCTCGAGC 23968
 QY 134 TAGCTCGAGCTCCCACTGCTCACTCTCTCTCCCACTGGGCGGCACTGCTAG 193

```

----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: hunquerry@sanger.ac.uk
Project Information
Center project name: bA697G3
Summary Statistics
-----
Assembly program: XGAP4, version 4.5
Sequencing vector: plasmid; T08752; 100% of reads

```

```
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
```

1. 169863

1. .169863

Derivative

AGCTGCAATCTG

GCA GCT C C C C A

GTCCTGCCAAG

EGG CATEGORIES

TTGCTGGCATG

GAA GCA GTG TG

CCALGGIGIG

TABLE 1

1
2
3
4
5
6
7
8
9
10
11
12

GGCTCTGGAGC

QY 734 ATGTCACCCGGGATG 749
 Db 24569 ATGTCACCCGGGATG 24584

RESULT 12
 AC114183/c
 LOCUS
 DEFINITION
 AC114183
 AC114183 207775 bp DNA linear HTG 13-MAY-2003
 Rattus norvegicus clone CH230-230F6, *** SEQUENCING IN PROGRESS
 *** 4 unordered pieces.
 AC114183
 HTG, HTGS_PHSSET; HTGS_DRAFT; HTGS_ENRICHED.
 Rattus norvegicus (Norway rat).
 SOURCE
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 207775)
 Muray, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alsbrooks, S., Amin, A., Angulano, D.,
 Anyabebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Crease, A., D'Souza, L.,
 Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Drepper, H., Dugan-Rocha, S., Dunn, A., Dublin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gebregorgis, Z., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
 Gurnatne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,
 Hollins, B., Howells, S., Hulik, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Kapathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kowals, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorensueta, L., Louisgeed, H., Lozano, R. J., Lu, X., Ma, J.,
 Mathew, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E.,
 Manthey, S., McLeod, M. P., McNeill, T. Z., Meenan, E.,
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwokolemehe, O., Okunolu, G., Olampunsaogon, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.,
 Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
 Sanders, W., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajda, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Sytek, A., Tabors, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K.,
 Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, A., White, F.,
 Williams, G., Willson, R., Wleczyk, R., Woodson, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, U., Yoon, L., Yoon, Y.,
 Yu, F., Zhang, J., Zhou, J., Zhou, S., Zhao, S., Dunn, D., von
 Niederhausen, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
 Weinstock, G. and Gibbs, R. A.
 Direct Submission

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 Rattus norvegicus
 Direct Submission
 Submitted (07-MAR-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 207775)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On May 13, 2003 this sequence version replaced gi:23605528.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GRP
 Center clone name: CH230-230F6
 ----- Summary Statistics
 Assembly program: Atlas 3.0;
 Consensus quality: 191020 bases at least Q40
 Consensus quality: 194390 bases at least Q30
 Consensus quality: 197174 bases at least Q20
 Estimated insert size: 203127; sum-of-contigs estimation
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 4 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 57140: contig of 57140 bp in length
 * 57141 57240: gap of unknown length
 * 57241 195381: contig of 138141 bp in length
 * 195382 195481: gap of unknown length
 * 195482 196790: contig of 1309 bp in length
 * 196791 196890: gap of unknown length
 * 196891 207775: contig of 10885 bp in length.
 Location/Qualifiers
 1. 207775
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-230F6"
 misc_feature
 223..1082
 /note="clone boundary
 clone end:17
 site:BCORI
 end sequence:B2112338"
 BASE COUNT 61791 a 35195 c 38071 g 59202 t 9516 others
 ORIGIN
 Query Match 24.4%; Score 432.4; DB 2; Length 207775;
 Best Local Similarity 90.1%; Pred. No. 1.1e-88;
 Matches 463; Conservative 0; Mismatches 51; Indels 0;

QY	236	GGTTAGGAACTGCGAGACGCGCGCCATGAG33CCGCCCTCGG3GCGGGGTCTCTGC	235
Db	172861	GGCCGGGCTCTGGCCCGAATGACACGCAATGAG33GCGGCTCCCGAGATCGGGGTCTACCTGC	172802
QY	296	CGCGGTGCTGCGGCTTTCTCAGATTGCTGAGATGATGCTTCCGTCGCGGACCGTAGTCCG	355
Db	172801	CGCGGCGGCTCGGAGCTGCGCGGCTGCTGAGCTTGGTGTCTTCTCTACCTGTGAGTCCG	172742
QY	356	CAGGCAACCGGTTCCCGGGGGGCTGAGCAGTGTGACACCGCGTACCTCAGCTGTGCTGG	415
Db	172741	CACGCGCCCGGTCTCCGGGGAGCGAAGCTGTGTGACACCGGTACTCAGCTGTCTGG	172682
QY	416	CGGTTTCGCGCACACGGAAGTGAACCGCTACGCTGTGAGACTGACGACGAGAGGCGCTGTAC	475
Db	172681	CGGTTTCGCGCACCCCGGAGTGAACCGCACGCTGTGTGAGACTGACGAGAGAGGCGCTGTAC	172622
QY	476	GGCGCAGACCTGCGCGCTGAGACCTGTGAGCTGTGGGGGCTCGGATACCGCGACGAGCGG	535
Db	172621	GGCGCAGACCTGCGCGCTGAGACCGCGGTCTCGGGGCTCTGTGTATCCGCGACGAGCGG	172582
QY	536	GGCGCTTAAACGCTGTGAACCCGACACGAATTCACACGCTGACACGTTTGGGAGAGACCC	595
Db	172581	GGCGCTCAACGCTGTGAACCCGACACGAATTCACAGGTGTACCTGTTTGGGAGAGACG	172502
QY	596	GTGCAAGTCTTGTGTTGGTGGCCCTCATTCACACGCGCGAGGCTGACCTTGCGAAGACAG	655
Db	172501	GTGCAAGATTTGTGTTGGTGGCCCTCATTCACACCGCGAGGCTGACCTTGCGAAGACAG	172442
QY	656	ATCCATCTGCGCTTATGAGAGATGAGGCGGTGTGAGAGCGCGTCACTTTAATCTCCCGGAGCC	715
Db	172441	ATCCATCTGCGCTTATGAGAGAGAGGCGCTTGTGAGAGCGCGTCACTTTAATCTCCCGGAGCC	172382
QY	716	CGCAATGAGTTCATCCCATGTCTACCCCGGGTG	749
Db	172381	CGCAATGAGTTCATTCCTCATGTCTACCCCGGGTG	172348
RESULT 13			
LOCUS	EX470094	252305 bp	DNA linear HTG 02-JUN-2003
DEFINITION	Mus musculus chromosome X clone RP23-130F16, *** SEQUENCING IN PROGRESS ***	32 unordered pieces.	
ACCESSION	EX470094	GI:31335952	
VERSION	EX470094.4	GI:31335952	
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 252305)		
AUTHORS	McClay, K.		
TITLE	Direct Submission		
JOURNAL	Submitted (30-MAY-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humbquer@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk		
COMMENT	On Jun 2, 2003 this sequence version replaced gi:30387055. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.		
	----- Genome Center		
	Center: Wellcome Trust Sanger Institute		
	Center code: SC		
	Web site: http://www.sanger.ac.uk		
	Contact: humbquer@sanger.ac.uk		
	----- Project Information		
	Center project name: BM130F16		
	----- Summary Statistics		
	Assembly program: XCAP4; version 4.5		
	Chemistry: Dye-terminator; 100% of reads		
	Consensus quality: 239922 bases at least Q40		
	Consensus quality: 243631 bases at least Q30		

Consensus quality: 245788 bases at least Q20	
Insert size: 249205; sum-of-contigs	
Insert size: 186861; 5.2% error; agarose-fp	
Quality coverage: 5.70x in Q20 bases; sum-of-contigs Quality	
coverage: 11.65x in Q20 bases; agarose-fp	

NOTES: This is a 'working draft' sequence. It currently	
* consists of 32 contigs. The true order of the pieces	
* is not known and their order in this sequence record is	
* arbitrary. Gaps between the contigs are represented as	
* runs of N, but the exact sizes of the gaps are unknown.	
* This record will be updated with the finished sequence	
* as soon as it is available and the accession number will	
* be preserved.	
1	40486: contig of 40486 bp in length
*	40487 40586: gap of 100 bp
*	40587 53971: contig of 13385 bp in length
*	53972 54071: gap of 100 bp
*	54072 65771: contig of 11700 bp in length
*	65772 65871: gap of 100 bp
*	65872 70703: contig of 4832 bp in length
*	70704 70803: gap of 100 bp
*	70804 72860: contig of 2057 bp in length
*	72861 72960: gap of 100 bp
*	72961 75036: contig of 2076 bp in length
*	75037 75136: gap of 100 bp
*	75137 77847: contig of 2711 bp in length
*	77848 77947: gap of 100 bp
*	77948 81334: contig of 3387 bp in length
*	81335 81435: gap of 100 bp
*	81435 83829: contig of 2395 bp in length
*	83830 83929: gap of 100 bp
*	83930 117462: contig of 33533 bp in length
*	83930 117462: gap of 100 bp
*	117463 117562: gap of 100 bp
*	117563 188799: contig of 71237 bp in length
*	188800 188809: gap of 100 bp
*	188900 198007: contig of 9108 bp in length
*	198008 198107: gap of 100 bp
*	198108 200705: contig of 2588 bp in length
*	200706 200805: gap of 100 bp
*	200806 202955: contig of 2147 bp in length
*	202955 203052: gap of 100 bp
*	203053 208517: contig of 5465 bp in length
*	208518 208617: gap of 100 bp
*	208618 211137: contig of 2520 bp in length
*	211138 211237: gap of 100 bp
*	211238 213763: contig of 2526 bp in length
*	213764 218663: gap of 100 bp
*	218664 216775: contig of 2812 bp in length
*	216776 216775: gap of 100 bp
*	219251 219251: contig of 2476 bp in length
*	219252 219351: gap of 100 bp
*	219352 221818: contig of 2467 bp in length
*	221819 221918: gap of 100 bp
*	221919 223941: contig of 2023 bp in length
*	223942 224041: gap of 100 bp
*	224042 226074: contig of 2033 bp in length
*	226075 226175: gap of 100 bp
*	226175 228553: contig of 2359 bp in length
*	228554 228633: gap of 100 bp
*	228634 231398: contig of 2765 bp in length
*	231399 231498: gap of 100 bp
*	231499 234893: contig of 3395 bp in length
*	234894 234993: gap of 100 bp
*	234994 237520: contig of 2527 bp in length
*	237521 237620: gap of 100 bp
*	237621 239638: contig of 2016 bp in length
*	239639 239735: gap of 100 bp
*	239737 241963: contig of 2227 bp in length
*	241964 242005: gap of 100 bp
*	242064 244300: contig of 2237 bp in length
*	244301 244400: gap of 100 bp
*	244401 246533: contig of 2139 bp in length

	misc_feature	237621..239636	/note="assembly_fragment:03414"	
	misc_feature	239737..241963	/note="assembly_fragment:03479"	
	misc_feature	242064..244300	/note="assembly_fragment:03553"	
	misc_feature	244401..246539	/note="assembly_fragment:04290"	
	misc_feature	246640..249593	/note="assembly_fragment:04434"	
	misc_feature	249694..252305	/note="assembly_fragment:04462"	
BASIS COUNT	73622 a	49859 c	48055 g	77640 t 3126 others
ORIGIN				
Query Match	24.2%, Score 429.2, DB 2, Length 252305;			
Best Local Similarity	91.4%; Prid. No. 6.2e-88;			
Matches 455,	Conservative	0;	Mismatches 43;	Indels 0; Gaps 0;
Dn	54543	GAGTGGCCGCCAATGAGGCGCCGCCGCCGAGATCGGGGTCTACCGCAGCGGC	54602	
Qy	312	TTCGCAATTGCTGCAACATGATGTTCTTGCGGCTTAAGTCCGACGACCAGTCCC	371	
Dn	54603	CTGCGCGCGTACTGAGCTTGAGCTTCTCTTAGCTCTGAATCGACAACGCGCCGATTCCC	54662	
Qy	372	GGGGGGCTAAACATGTGTGAACCGCGAACCAACAGTCTCGAGCGGGTTCGACACG	431	
Dn	54663	GGGAGACGAAACCGGTGTGAACCGCGAACCAACAGTCTCGAGCGGGTTCGACACG	54722	
Qy	432	GAGTGAACCGTACGAGTGTGAGAGCTGAGCGAGAGGCGCGTGAACGACAGATCGCGG	491	
Dn	54723	GAGTGAACCGTACGAGTGTGAGAGCTGAGCGAGAGGCGCGTGAACGACAGATCGCGG	54782	
Qy	492	TGGAAGCTGTGGCTTGGGGTCTCTGTATACGCCCCAGCGCGCGCGGCTTAAAGCTGTAT	551	
Dn	54783	TGGAAGCTGTGGCTTGGGGTCTCTGTATACGCCCCAGCGCGCGCGGCTTAAAGCTGTAT	54842	
Qy	552	ACCGGACACGAAATTCAACGTATCCCAACGAGTTGGGGAGACACCGTGAAGTCTGTGAT	611	
Dn	54843	ACCGGACACGAAATTCAACGTATCCCAACGAGTTGGGGAGACACCGTGAAGTCTGTGAT	54902	
Qy	612	TGCGCCTCATCCAAACGCGCGCGGGGCTGCACCTTGCGACACAAAGATCATCTGGCTATG	671	
Dn	54903	TGCGCCTCATCCAAACGCGCGGGGCTGCACCTTGCGACACAAAGATCATCTGGCTATG	54962	
Qy	672	AAGAATGGGGGTCTGAAACCGGTCACTTTTAACTTCCCGGAGACCGGCAATGAGGTATCC	731	
Dn	54963	AAGAATGGGGGTCTGAAACCGGTCACTTTTAACTTCCCGGAGACCGGCAATGAGGTATCC	55022	
Qy	732	CCATGCTCACCGCGGGTGTG	749	
Dn	55023	CCATGCTCACCGCGGGTGTG	55040	
RESULT 14				
LOCUS	AC109686	233385 bp	DNA linear	HTG 11-OCT-2002
DEFINITION	Rattus norvegicus clone CH230-231D21, *** SEQUENCING IN PROGRESS			
ACCESSION	AC109686			
KEYWORDS	HTG, HTGS_PHASE1, HTGS_DRAAF, HTGS_ENRICHED.			
SOURCE	Rattus norvegicus (Norway rat)			
ORGANISM	Rattus norvegicus			
REFERENCE	1 (bases 1 to 233985)			
AUTHORS	Murray D.Marie, Merker,M.Isee, Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alstbrooks,S., Amin,A., Anguiano,D., Anyalbech,V., Ayvagi,A., Ayodeji,W., Bada,B., Baden,H.,			

QY	356	AAGCAACCGCGTTCCTCCGGGGGGGCTGAAGAGATGGTGAACGCGCATCTTCAACGGTCTCGG	415
Db	228442	CACGCGCCGGTTCCCGGGAGCCGAGCTGTGTGAACCGCGTACTTCAACGTGCTCGG	228501
QY	416	CGGGTTCCGACACGGAGATGAACCGTACGGGTGTGGAGAGCTAGAGAGAGCGCGTATC	475
Db	228502	CGGGTTCCGACCCGAGATGAACCGACGGGTGTGGAGAGCTAGAGAGAGCGCGTATC	228561
QY	476	GGCCAGGACTTCGCGCGCTGAGACCTGT-GGCTGGGGTCTTGTATCCGCCCAAGGACCGG	534
Db	228562	GGCCAGGACTTCGCGCGCTGAGACCGGTCTTCGGGGGGCGCTGTATCCGCCCAAGGACCGG	228621
QY	535	GGGCGCTTAAGCGCTGTAAACCGGACACGAATTTACGGGTGGCCACGGTTGGGGAGAGAC	594
Db	228622	GGCGCTTAAGCGCTGTGAACCGGACACGAATTTACGGGTGGCCACTGTTGGGGAGAGAC	228681
QY	595	CGTGCAAGTCTCTGTATTGGCCCTCATCAACGGCGGGGGCTGCATCTTCCACACAA	654
Db	228682	GGTGCAAGTCTCTGTATTGGCCCTCATCAACGGGGGGGTGCATCTTCCGGACAA	228741
QY	655	GATCATCTGGCTTATGAGAGATGGGGCTTGGAGCCGTATCTTTAACTTCCCGGGAC	714
Db	228742	GATCATCTGGCTTATGAGAGAGGGGCTTGGAGCCGTATCTTTAACTTCCCGGGAC	228801
QY	715	CGGCAATGAGGTCAACCCCAATCTCAACCGGGTG	749
Db	228802	CGGCAATGAGGTCAATCCCAATCTCAACCCGGTG	228836

Db	187	GAAAGATGAAACGGCTCTGGAGGAAACAGTGAATCAAAATGAAAGCTCAAGCTGG	246
QY	1434	TAAACCATGAACAAATTCCTGGCAGTGAATCTTATCTCAAGTGGAAACCCAACT	1493
Db	247	TAAACCATGAAGCAATTCCTGGCAGTGAATCTTATCTCAAGTGGAAACCCAACT	306
QY	1434	TTGGAAGAGACAAACTCTTATCAAGAGCTGCTGTTGAGAAATTTAAATCTTAAATTC	1553
Db	307	TTGGAAGAGACAAACTCTTATCAAGAGCTGCTGTTGAGAAATTTAAATCTTAAATTC	366
QY	1554	TGTGTAATGAAAACCTGAACCATGTGATTAACGA	1591
Db	367	TGTGTAATGAAAACCTGAACCATGTGATTAACGA	404

Search completed: October 6, 2003, 16:32:49
 Job time : 6529 secs

```
Search completed: October 6, 2003, 16:32:49
Job time : 6529 secs
```

LOCUS	AX071980	404 bp	DNA	linear	PAT 25-JAN-2001
DEFINITION	Sequence 2452 from Patent WO0102568.				
ACCESSION	AX071980				
VERSION	AX071980.1	GI:12582331			
KEYWORDS	.				
SOURCE	Homio sapiens (human)				
ORGANISM	Homio sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butiheria; Primates; Catarrhini; Homiidae; Homo.				
AUTHORS	1 Williams, L.T., Escobedo, J., Imis, M.A., Garcia, P.D., Klinger, J., Kassam, A., Reinhard, C., Randazzo, F., Kennedy, G.C., Pot, D., Lamson, G., Demian, R., Crkeniskov, R., Drmanac, S., Dickson, M., Labat, I., Ieshchowitz, D., Kite, P., Garcia, V. and Strache-Cran, B.				
TITLE	Human genes and gene expression products				
JOURNAL	Patent: WO 0102568-A 2452 11-JAN-2001; CHIRON CORPORATION (US); HVSQ, INC. (US)				
FEATURES	Location/Qualifiers				
SOURCE	1..404				
	/organism="Homio sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
BASE COUNT	134 a 81 g 90 g 59 t				
ORIGIN					

	Query Match Best Local Similarity Matches 39;	22.1%; 99.0%; Conservative 0;	Score 392.2; Pred. No. 5.2e-80; Mismatches 4;	DB 6; Length 404; Indels 0;	Caps 0;
QY	1194	ACAAGACTGCCCATATGCAATATGACATATCAAGCTTTGGAAATGAGGTGATG	1255		
DB	7	AGAGGACTGCCCATATGCAATATGACATATCAAGCTTTGGAAATGAGGTGATG	66		
QY	1254	TGGAATATGATCAGTGTCTTTCAAGTCCCTGTATTCATGAAATATCTAATATGTCCT	1311		
DB	67	TGGAAGATGATCAGTGTCTTTCAAGTCCCTGTATTCATGAAATATCTAATATGTCCT	126		
QY	1314	CTCCCATGAGAGAGATATCGACGACGACCCGATCATATGCAATATCTTCACATACAGG	1372		
DB	127	CTCCCATGAGAGAGATATCGACGACGACCCGATCATATGCAATATCTTCACATACAGG	186		
QY	1374	GAACGATGAACCCGCTCTCGAGAGAACGTGACACTCAACAAATGAAGCTTCAACGCTGG	1433		

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 6, 2003, 14:32:48 ; Search time 3882 Seconds

(without alignments)
11106.682 Million cell updates/sec

Title: US-09-854-300-7

Perfect score: 1774

Sequence: 1 agctgcagctccacgcgcgt.....atagcgaagttccctcag 1774

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 1215238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estdb:*
2: em_esthm:*
3: em_estin:*
4: em_estim:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estum:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrt:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1205.2	67.9	1921	AK018582	AK018582 Mus muscu
2	1202	67.8	2741	AK004847	AK004847 Mus muscu
3	1003	56.5	2677	BC036301	BC036301 Homo sapi
4	847.6	47.8	1201	EX343985	EX343985 EX343985

5	836.8	47.2	1954	11	BC030951	BC030951 Homo sapi
6	746.4	42.1	861	14	CD107148	CD107148 AGNCOURT
7	680.4	38.4	1806	11	BC012931	BC012931 Homo sapi
8	673.8	38.0	811	12	B1600046	B1600046 603251431
9	654.2	36.9	1000	12	B1600419	B1600419 603246458
10	644.2	36.3	676	9	AV690200	AV690200 AV690200
11	641	36.1	773	12	B1601096	B1601096 603249588
12	623.4	35.1	840	11	AK008312	AK008312 Mus muscu
13	622.8	35.1	1150	12	B1488809	B1488809 603021137
14	622.4	35.1	906	10	B1615370	B1615370 602346005
15	622.2	35.1	683	9	AV686169	AV686169 AV686169
16	615.4	34.7	996	10	B1476206	B1476206 602333351
17	609.8	34.4	866	14	CB195525	CB195525 AGNCOURT
18	577	32.5	681	14	B1220592	B1220592 602935079
19	576.6	32.5	832	14	B1708243	B1708243 B1708243
20	568	32.0	1001	14	B1705053	B1705053 B1705053
21	564.4	31.8	683	9	AV692456	AV692456 AV692456
22	557	31.4	965	14	CB235518	CB235518 AGNCOURT
23	548	30.9	835	13	B1601345	B1601345 AGNCOURT
24	542.8	30.6	557	12	BMS07037	BMS07037 h24805.y
25	537.2	30.3	904	10	B1664617	B1664617 602117901
26	531.2	29.9	596	10	B1231341	B1231341 253699.BA
27	524.2	29.5	862	10	B1668554	B1668554 602123601
28	520.8	29.4	756	14	CB389957	CB389957 AGNCOURT
29	515.6	29.1	902	10	B165006	B165006 602343868
30	513.8	29.0	841	10	B1696359	B1696359 602124911
31	509.8	28.7	812	12	B1110345	B1110345 602900777
32	502.2	28.3	872	10	B1699229	B1699229 602126878
33	499.4	28.2	789	10	B1669635	B1669635 602120364
34	496	28.0	759	10	B1666897	B1666897 602121530
35	492.4	27.8	759	10	B1640079	B1640079 602146572
36	482.4	27.2	774	10	B16432146	B16432146 602496525
37	481.6	27.1	841	10	B1666680	B1666680 602121368
38	481.6	27.1	847	10	B1658515	B1658515 601645138
39	475.6	26.8	854	12	B1653069	B1653069 603300779
40	471.8	26.6	856	10	B1658087	B1658087 601644929
41	471.2	26.6	717	13	B101982	B101982 BX101982
42	469	26.4	680	10	B1609809	B1609809 601095109
43	468.6	26.4	884	10	B1642034	B1642034 602063189
44	466.4	26.3	835	10	B1658350	B1658350 601644838
45	461.4	26.0	622	10	B137973	B137973 601784651

ALIGNMENTS

RESULT 1
AK018582
LOCUS
DEFINITION Mus musculus adult male cecum cDNA, RIKEN full-length enriched library, clone:130001F19 product:Grail homolog [Homo sapiens], full insert sequence.
ACCESSION AK018582
VERSION AK018582.1 GI:12858361
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL MEDLINE PUBMED
20499374 11042159

QY	672	AGAGATGGGGGTGTGAGACCGCATCTTAACTCCCGSAGACCGGAAAGAGATCAC	731
Db	598	AGAGAGGGGCTTTGAGGGGTCACTTTTACTCTCCGAGACCCGAAATGAGTCAACC	657
QY	733	CCATGTCTCACCCGGGTGCAATGACATGTTGCAATCATGATCGCAATCTGAAGACA	791
Db	658	CCATGTCTCACCGGGTGTGGGGCAATGTTGCATCATGATTTGGCAATCTGAAGACA	717
QY	792	CAAAATTTGCAATCTTATCTAAAGAGCATACAGTGCATATGTCATAGAAATGAGGA	851
Db	718	CAAAATTTGCAATCTTATCTAAAGAGCATACAGTGCATATGTCATAGAAATGAGGA	777
QY	852	AAAAACATGACCCCTTGGGATGACATCTTCAATTTTGTGTGTGTGCTCTTTTAA	911
Db	778	AAAAACATGACCCCTTGGGATGACATCTTCAATTTTGTGTGTGTGCTCTTTTCA	837
QY	912	TTATTAAGGGGGAACCTGTGGCTATTTTATCTTTTATCTGTCTGAGAGCTACGAAAG	971
Db	838	TTATTAAGGGGGAACCCGTGGCTATTTTATCTTTTATCTGTCTGAGAGATACGAATG	897
QY	972	CAAGAGCTCAACACAGAAAGACAGAGGCATTAAAGSAGATGCTAAAAAGCTATGGA	1031
Db	898	CAAGAGCTCAACACAGAAAGACAGAGGCATTAAAGSAGATGCTAAAAAGCTATGGA	957
QY	1032	GGCTTCAATACGACACTGAACAAGAGACAAGSAAATGGCCCTGATGAGATGTT	1091
Db	958	AGCTTCAAGCTCCCACTTTGAACAAGAGACAAGSAAATGGCCCTGATGAGATGCT	1017
QY	1092	GTGGCTGTGCAATGAAATGTGTAAACCAATATATTGGTACGATCTPAAGTCGACAC	1151
Db	1018	GTGGCTGTGTGATGAGCTCTATTAAGCCAAATGATTTGGTGGCACTCCATCTGCATC	1077
QY	1152	AATATTTCCAAAGACATGTGTTGACCATGCTGTATTAACAAGACATTTGCCCATG	1211
Db	1078	AATATTTCCAAAGACATGTGTTGACCATGCTGTATTAACAAGAGCTTTGCCCATG	1137
QY	1212	GCATATGTGACATCTCAAAAGCTTTGGGAATGAGGTGAAATTTGAAGATGATCGAT	1271
Db	1138	GCATATGTGACATCTCAAAAGCTCTGGGAATGAGGTGAAATTTGAAGATGATCGAT	1197
QY	1272	CTTTCAAGTCCCGTATCCATGAAATCTAATAGTGCCTCCCTCCCATGAGAGATTA	1331
Db	1198	CTTTCAAGTCCCGTATCTAATGAGATCTAATAGTGCCTCCCTCCCATGAGAGATTA	1257
QY	1332	ATCGACGAGACCGCATATTTGGATATGCTTCAAGTACAGAGAAACAGATGAACGCTC	1391
Db	1258	ATCGACGAGACCGCATATCTGATATGCTTCAAGTACAGAGAAACAGATGAACGCTC	1317
QY	1392	TGGAGGAAACATGACAGTCAAAATGSAATCTCAACGCTGTAAACATGAGACAAAT	1451
Db	1318	TGGAGGAAACATGACAGTCAAAATGSAATCTCAACGCTGTAAACATGAGACAAAT	1377
QY	1452	CTGTGGCAGTGAATGTATTTCTCATGTGACAAACCAACCTTTGAAGAGAGAAATCT	1511
Db	1378	CTGTGGCAGTGAATGTATTTCTCCCATGTGACAAACCAACCTTTGAAGAGAGAAATCT	1437
QY	1512	CTTAACAGAGACCTGCTGTGAGAGAAATTAATCT--AAATCTGTGTAAATGAGAAAT	1570
Db	1438	CTTAACAGAGAGCAGCTGTGAGAGAAATTAATCTTAATTAATCTGTGTGTGTAAATCT	1497
QY	1571	TGACCATTTAG--TATTAACAGACCTGCAATCAAGGCTAGATT--CTATTAATTAATTG	1628
Db	1498	TGACCATTTAGTTTAACAAAGAGCTGCCAATCAAGGCTAGATTCTATGAATGAACCTG	1557
QY	1629	ATTAATTTAATTAATTAAGATGATCTGAAGAGTCTCAAGATGATTAATTAATGCTATA	1688
Db	1558	ATTAATTTAATTAATTAAGATGATCTGAAGAGTCTCAAGATGATTAATTAATGCTATA	1614
QY	1689	GTTAATAATGGCTTAAATATTTAACTGTATACTTTTTCACAAATCATTAATATA	1748
Db	1615	GTTAATA--TGCTTAATTAATTTAATCCCAAGTACGTTTT--CCAAATCTACCAATAG	1670
QY	1749	TTTTTCATAGGCAAGTTCTCTCT	1771

Db	1671	TTTTTCATAGGCAAGTTTCCTCT	1673
RESULT 2			
AK004847		2741 bp	mRNA
LOCUS			linear
DEFINITION			musculus adult male liver cDNA, RIKEN full-length enriched library, clone:1300002C13 product:SRAL1 homolog [Homo sapiens], full insert sequence.
ACCESSION	AK004847		
VERSION	AK004847.1	GI:12836339	
KEYWORDS	HTC; CAP trapper		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE			
AUTHORS	Carninci, P. and Hayashizaki, Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PIRME	10349636		
REFERENCE			
AUTHORS	2 Carninci, P., Shibata, Y., Hayatsu, N., Sugihara, Y., Shibata, K., Itoh, M., Komoto, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
MEDLINE	20499374		
PIRME	11042159		
REFERENCE			
AUTHORS	3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komoto, H., Akiyama, J., Nishi, K., Kitsunai, T., Yashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wachihi, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multichannel sequencer		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		
MEDLINE	20530913		
PIRME	11076861		
REFERENCE			
AUTHORS	4 Kawai, J., Shingawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, T., Hara, A., Fukunishi, Y., Komoto, H., Aichi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavola, T., Fleischmann, W., Gasteirland, T., Gissi, C., King, B., Kochwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schiraldi, L. M., Stambli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, I., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Botfield, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Butt, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Nombarts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, T., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyao-oka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L., Wysocki-Borja, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S. and Hayashizaki, Y.		
TITLE	Functional annotation of a full-length mouse cDNA collection		
JOURNAL	Nature 409 (6821), 685-690 (2001)		
MEDLINE	21085660		
PIRME	11217851		
REFERENCE			
AUTHORS	5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.		
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs		

JOURNAL
REFERENCE
AUTHORS

Nature 420, 563-573 (2002)

Adachi, J., Aizawa, K., Akhita, S., Akiyama, T., Arai, A., Aono, H.,
Arakawa, T., Bono, H., Carlini, P., Fukuda, S., Fukunishi, Y.,
Furuno, M., Hanagaki, T., Haru, A., Hayatsu, N., Hiramoto, K.,
Hirooka, T., Hori, F., Imoto, K., Ishii, Y., Itoh, M., Iwase, M.,
Kasukawa, T., Kato, H., Kawal, J., Kojima, Y., Komori, H., Konda, M.,
Koyu, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishik, K.,
Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
Oshiba, K., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
Suzuki, H., Tagami, M., Tasawa, A., Takehashi, F., Tanaka, T.,
Tejima, Y., Toyota, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
Yoshino, M., Mizumatsu, M. and Hayashizaki, Y.

TITLES
JOURNAL

Submitted 10-07-2009 Yoshinori Hayashiyaiki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Saito-cho, Tsukuba-Ku, Yokohama Kanagawa 230-0045, Japan (E-mail: yoshinori@gscc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

FEATURES

SOURCE

cDNA library was prepared and sequenced in mouse genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer 15', GAGAGAGAGAGCGCCGACATCGAGTTTCTTTTCTTTTATN 3'. cDNA was prepared by using Trehalase thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence 15' GAGAGAGAGAGGATCGAGACCTTATTTATTTATTAACCCCCCCC 3'. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end: XhoI. Host: SOLA.

CDS

```

cds
/codon_lib="Riken full-length enriched mouse cDNA library"
/def_stages="adult"
188..1474
/notes="unnamed protein product; GRAL1, homolog [Homo sapiens] (Spur, 0968F7, evidence: FASSTV, 93.2%ID, 100%length, match=1284) putative"
/codon_start=1
/protein_id="PAB23613.1"
/db_xref="GI:12836340"
/db_xref="MG1:1914139"
/translators="MGPPPGICVYGRGCGAGARLLAMCFLLASLSPHAPQSRGAENVWTAALNLSVLTGQVNRVTWELLESSEYVGQDSELPYSGVLPDGPQALNCPNNPNNIVPYWVSCTVSWLALIORGASCTPADNNLTHLSEKSGASGAVIPNPGRNVSFMSHLPAGVIAIMTGNMKGTKTILOSIORGIQWMTVIEKKGKGVWVHSIPFVSVSIFLITQATGVYIPASARLIRNARASQSRKQSLAKDKAKLQKLTQKQKGTICPGPQSSCAVCEIYKIPMDLIRLITGNHIFKTCVDPMDLHKKRQKQKLTALGISEVDPGQSSSYSLQVPSVMSAENYSLSPHEEDSRSTASSGASVQADDEPPLBEAQSANRLQLVANSEANVADVDPHNDPTEDEETDPEAAVREIKS"
2725..2730
polyA signal

```

polyA_site

```
/note="putative"
2741
```

BASE COUN
ORIGIN

764 a 578 c 643 g 756 t

[illegible]

RESULT 3				
LOCUS	BC036901			
DEFINITION	BC036901	2677 bp	mRNA	linear
ACCESSION	BC036901		Homo sapiens, clone IMAGE:553474,	HTC 04-MAR-2003
VERSION	BC036901.1			
KEYWORDS	HTC.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia, Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
TITLE	1 (bases 1 to 2677)			
JOURNAL	Strausberg, R.			
	Direct Submission			
	Submitted (23-AUG-2002) National Institutes of Health, Mammalian			
	Gene Collection (MGC), Cancer Genomics Office, National Cancer			
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,			
	USA			
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov			
COMMENT	Contact: MGC help desk			
	Email: cgapps-d@mail.nih.gov			
	Tissue Procurement: ATCC			
	cDNA Library Preparation: Life Technologies, Inc.			
	cDNA library Arrayed by: The I.M.A.G.E. Consortium (ILMN).			
	DNA Sequencing by: National Institutes of Health Intramural			
	Sequencing Center (NISC),			
	Gaithersburg, Maryland;			
	Web site: http://www.nisc.nih.gov/			
	Contact: nisc.mc@hgri.nih.gov			
	Akheri,N., Ayele,M.C., Beckstrom-Stenberg,S.M., Benjamin,B.,			
	Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,			
	Dierich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,			
	Ramsen,N., Ho,S.-U., Karlins,S., Kwong,P., Latic,F., Legaspi,R.,			
	Maduro,Q.L., Mastello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,			
	McDowell,T., Pearson,R., Stantiripop,S., Thomas,P.J., Touchman,J.W.,			
	Tsugeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,D.,			
	Young,A., Zhang,L.-H. and Green,E.D.			

1 253

QY 1417 TGAAGTCTACAGCTGGTAAACCATGAGCAATTCGTGGCAGTGGATGTATTCTCA 1476
 Db 1254 TGAAGTCTACAGCTGGTAAACCATGAGCAATTCGTGGCAGTGGATGTATTCTCA 1313
 QY 1477 TGTGACCAACCACTTGAAGAGACCACTCTATCAAGAGTCTGTTCGAGA 1536
 Db 1314 TGTGACCAACCACTTGAAGAGACCACTCTATCAAGAGTCTGTTCGAGA 1373
 QY 1537 AATTAAATCTTAAATCTGTGTAAATGAAACCTTGAACCTTATGAAATCAAGACTGC 1596
 Db 1374 AATTAAATCTTAAATCTGTGTAAATGAAACCTTGAACCTTATGAAATCAAGACTGC 1433
 QY 1597 CAATCAGGCGCTAGTTCCTATATATATATGTAATTAATTAATTAATTAAGTATCT 1656
 Db 1434 CAATCAGGCGCTAGTTCCTATATATATGTAATTAATTAATTAATTAAGTATCT 1493
 QY 1657 GAAAGTCTCAGATGACTAATATATATGTAATTAATTAATTAATTAATTAATTAAT 1716
 Db 1494 GAAAGTCTCAGATGACTAATATATATGTAATTAATTAATTAATTAATTAATTAAT 1551
 QY 1717 GTTAACCTTTTCCACCAACCTGATATATATATTTTTCATAGGCAAGTTCCTCGAG 1774
 Db 1552 GTTAACCTTTTTCACCAACCTGATATATATATTTTTCATAGGCAAGTTCCTCGAG 1608
 RESULT 4
 EX343985 1201 bp mRNA linear EST 01-MAY-2003
 LOCUS BX343985 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 DEFINITION clone CS0D1043YB06 5-PRIME, mRNA sequence.
 ACCESSION BX343985
 VERSION BX343985.1 GI:30309065
 KEYWORDS EST.
 SOURCE
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1201)
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 COMMENT
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 11075.f For
 more information about this cluster, see
 http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS1A1011ZF040P1
 &cluster=11075.f. Contact : Peng Liang Email : filiang@life.com
 URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope, sequence ID : CS1A1011ZF040P1.
 FEATURES
 source
 1..1201
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1043YB06"
 /issue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="First strand cDNA was primed with a NotI-oligo (dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."
 BASE COUNT 183 a 323 c 344 g 258 t 93 others
 ORIGIN
 Query Match 47.8%; Score 847.6; DB 13; Length 1201;
 Best, Local Similarity 89.7%; Pred. No.1.2e-165;
 Matches 931; Conservative 46; Mismatches 52; Indels 9; Gaps 7;
 74 CCAGAGAGCTGCATCTGCGCAACCTGTGTGACGCTACGCTCTCTGCTCCAG 133

Db 75 CGAGAGAGCTGCATCTGCGCAACCTGTGTGACGCTACGCTCTCTGCTCCAG 134
 QY 134 TACCTGCAAGCTCCCAAGTCTCACTCAATTCCTTCCCACTGCGCACTGCTCAAG 193
 Db 135 TACCTGCAAGCTCCCAAGTCTCACTCAATTCCTTCCCACTGCGCACTGCTCAAG 194
 QY 194 ACCAGGCTCCCAAGTCTCACTCAATTCCTTCCCACTGCGCACTGCTCAAG 253
 Db 195 ACCAGGCTCCCAAGTCTCACTCAATTCCTTCCCACTGCGCACTGCTCAAG 254
 QY 254 GCGCGCGCCATGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 312
 Db 255 GCGCGCGCCATGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 314
 QY 313 TTCCCAATTCCTGCGCAAGTCTCTCTGCTGCGCGCGCGCGCGCGCGCGCGCG 372
 Db 315 TTCCCAATTCCTGCGCAAGTCTCTCTGCTGCGCGCGCGCGCGCGCGCGCGCG 374
 QY 373 GGGGCGTGAAGAGTGTGACCCGCACTCACTGCTGCGCGCGCGCGCGCGCGCGCG 432
 Db 375 GGGGCGTGAAGAGTGTGACCCGCACTCACTGCTGCGCGCGCGCGCGCGCGCGCG 434
 QY 433 AGTGAACCGTACGCTGTGCGAGCTGAGCGAGAGGCGCTGTACGCGCAGACTCCCGCT 492
 Db 435 AGTGAACCGTACGCTGTGCGAGCTGAGCGAGAGGCGCTGTACGCGCAGACTCCCGCT 494
 QY 493 GGAAGCTGTGAGGCGCTGTGCGAGCTGAGCGAGAGGCGCTGTACGCGCAGACTCCCGCT 552
 Db 495 GGAAGCTGTGAGGCGCTGTGCGAGCTGAGCGAGAGGCGCTGTACGCGCAGACTCCCGCT 554
 QY 553 CCGGCAACGAATTTCAAGTGTGCGAGCTGAGCGAGAGGCGCTGTACGCGCAGACTCCCGCT 612
 Db 555 CCGGCAACGAATTTCAAGTGTGCGAGCTGAGCGAGAGGCGCTGTACGCGCAGACTCCCGCT 614
 QY 613 GGCCTCATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 671
 Db 615 GGCCTCATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 674
 QY 672 AGAAGTGTGCGCTGTGAGAGCGCTGTACCTTCCCGGAGCGCGCAATGAGTCAATCC 731
 Db 675 AGAAGTGTGCGCTGTGAGAGCGCTGTACCTTCCCGGAGCGCGCAATGAGTCAATCC 734
 QY 732 CCATGCTCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 791
 Db 735 CCATGCTCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 793
 QY 792 CAAAAATTCGCAATCTATTCAAAGAGCAATGAGCAATGAGTCAATGAGAGGGA 851
 Db 794 CAAAAATTCGCAATCTATTCAAAGAGCAATGAGCAATGAGTCAATGAGAGGGA 852
 QY 852 AAAAAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 911
 Db 853 AAAAAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 912
 QY 912 TTATTAAGCGCGCACTGTGCGCAATTCCTTATTCCTTATTCCTGCGAGAGTCAAGGAT 970
 Db 913 TTATTAAGCGCGCACTGTGCGCAATTCCTTATTCCTTATTCCTGCGAGAGTCAAGGAT 972
 QY 971 GCAAGGCTCAAG 1030
 Db 973 GCAAGGCTCAAG 1031
 QY 1031 AGGCTTCACTGAGCACTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1090
 Db 1032 AGGCTTCACTGAGCACTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1088
 QY 1091 TGTGCTGTGCTGCTGAA 1108
 Db 1089 GYKTKGCTTGAATTTTAA 1106
 RESULT 5

BC030951
 LOCUS BC030951 1954 bp mRNA linear HTC 13-JUN-2002
 DEFINITION Homo sapiens, hypothetical protein FLJ23516, clone IMAGE:4280352,
 mRNA.
 ACCESSION BC030951
 VERSION BC030951.1 GI:21410659
 KEYWORDS HTC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 1954)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (03-JUN-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,
 USA
 REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 COMMENT Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC
 CNA Library Preparation: CLONTECH Laboratories, Inc.
 CNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www.shgc.stanford.edu>
 Contact: (Dickson, Mark) medpaxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>
 Series: IRAL Plate: 41 Row: C Column: 10
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 13375696
 This clone has the following problem: frame shifted.
 FEATURES
 source
 1..1954
 /location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="locusid:79589"
 /db_xref="taxon:9606"
 /clone="IMAGE:4280352"
 /rissue_type="Brain, Primitive neuroectodermal"
 /clone_lib="NIH_MGC_56"
 /lab_host="DH10B"
 /note="vector: pDNR-LIB"
 BASE COUNT 676 a 294 c 346 g 638 t
 ORIGIN
 Query Match 47.2%; Score 836.8; DB 11; Length 1954;
 Best Local Similarity 98.7%; Pred. No. 2.2e-163; Indels 3; Gaps 2;
 Matches 864; Conservative 0; Mismatches 8;
 QY 900 TGTCTTTTATATTAAGCGGCACTGTGGCTATTATCTTTATCTGTGCGAA 959
 Db 1 TGTCTTTTATATTAAGCGGCACTGTGGCTATTATCTTTATCTGTGCGAA 60
 QY 960 GGTACGGATGAGAGCTCAAGCGAGGAGAGGCAATTAAGCGAGATGCTAA 1019
 Db 61 GGTACGGATGAGAGCTCAAGCGAGGAGGCAATTAAGCGAGATGCTAA 120
 QY 1020 AAGCTATTGGAAGGCTCACTAGCACTGAAACAAGAGACAGAAATTTGCGCTG 1079
 Db 121 AAGCTATTGGAAGGCTCACTAGCACTGAAACAAGAGACAGAAATTTGCGCTG 180
 QY 1080 ATGAGATTAAGTGTGTGTGTGCAATGATGATTAACCAAGATTTGTGACGACT 1139
 Db 181 ATGAGATTAAGTGTGTGTGTGCAATGATGATTAACCAAGATTTGTGACGACT 240
 QY 1140 TAACTGCAACATATTTCCATTAAGCATGTGTGACCGATGCTGTAAACAAGA 1199

Db 241 TAACTGCAACATATTTCCATTAAGCATGTGTGACCGATGCTGTAAACAAGA 300
 QY 1200 CTTCGCCCATGTGCATATGTCATCTCTAAAGCTTTGGGAATAGGTGATTTGAAG 1259
 Db 301 CTTCGCCCATGTGCATATGTCATCTCTAAAGCTTTGGGAATAGGTGATTTGAAG 360
 QY 1260 AAGATCAAGTCTCTCAAGTCCCTGATCCAAATGAAATATATAGTGCCTCTCC 1319
 Db 361 AAGATCAAGTCTCTCAAGTCCCTGATCCCAATGAAATATATAGTGCCTCTCC 420
 QY 1320 ATGAAGAGATATGCGAGGAGACCGCATCATCTGTGATGCTTCAATGAGGAGAG 1379
 Db 421 ATGAAGAGATATGCGAGGAGACCGCATCATCTGTGATGCTTCAATGAGGAGAG 480
 QY 1380 ATGAACCGCTCTGAGGAGACCGCATCATCTGTGATGCTTCAATGAGGAGAG 1439
 Db 481 ATGAACCGCTCTGAGGAGACCGCATCATCTGTGATGCTTCAATGAGGAGAG 540
 QY 1440 ATGAACCAATCTGTGCGAGTGAATGTTATCTCATGTGACAAACCACTTTGAAG 1499
 Db 541 ATGAACCAATCTGTGCGAGTGAATGTTATCTCATGTGACAAACCACTTTGAAG 600
 QY 1500 AAGAGAACTCTTAATCAAGACCTGCTTCAAGAAATTAATCTTAATCTGTGA 1559
 Db 601 AAGAGAACTCTTAATCAAGACCTGCTTCAAGAAATTAATCTTAATCTGTGA 660
 QY 1560 AATGAAATCTGAACCAATTAATCAAGACCTGCTTCAAGAAATTAATCTTAAT 1619
 Db 661 AATGAAATCTGAACCAATTAATCAAGACCTGCTTCAAGAAATTAATCTTAAT 720
 QY 1620 AATGAAATCTGAACCAATTAATCAAGACCTGCTTCAAGAAATTAATCTTAAT 1679
 Db 721 AATGAAATCTGAACCAATTAATCAAGACCTGCTTCAAGAAATTAATCTTAAT 780
 QY 1680 TATGCTATGATTAATGCTTAAATATTTTAACTGTTTCACTTTTCACTTCA 1739
 Db 781 TATGCTATGATTAATGCTTAAATATTTTAACTGTTTCACTTTTCACTTCA 837
 QY 1740 AATTAATATTTTTCATGAGGAGTTCTCTCAG 1774
 Db 838 AATTAATATTTTTCATGAGGAGTTCTCTCAG 872
 RESULT 5
 LOCUS CD107148 861 bp mRNA linear EST 15-MAY-2003
 DEFINITION AGSCNOURT 14019634 NIH MGC 179 Homo sapiens cDNA clone
 IMAGE:30366694 5', mRNA sequence.
 ACCESSION CD107148
 VERSION CD107148.1 GI:30760322
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 861)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Dr. Michael Brownstein
 CNA Library Preparation: Invitrogen Corp
 CNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LIML at:
<http://image.llnl.gov>
 Plate: ND9426 Row: k Column: 23
 High quality sequence stop: 739.
 Location/Qualifiers
 1..861
 /organism="Homo sapiens"
 source

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30366594"
/tissue_type="Pituitary"
/lab_host="DH10B-Ton A ( T1 and T5 phage resistances )"
/clone_1b="NH_MGC_179"
/note="Organ: brain; Vector: pCMV-Sport6.1; Site_1: EcoRV
(destroyed); Site_2: NotI; Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon cloning
). Average insert size 1.1 kb. Library was constructed by
(Invitrogen). Note: this is a NH_MGC Library."

```

Query Match	42.1%	Score	746.4	DB 14	Length	861			
Best Local Similarity	96.4%	Ped. No.	1.3e-144						
Matches	785	Conservative	0	Mismatches	24	Indels	5	Gaps	2
Qy	963	TACGAAATGCAAGGCTCCAAACACAGAAACGAGGCAATTAAAGCGAGTGTCTAAAAAG	10222						
Db	1	TACGAAATGCAAGGCTCCAAACACAGAAAGAGGGCAATTAAAGCGAGTGTCTAAAAAG	60						
Qy	1023	CTATTGGAAGGCTCAACTACGCACTCTGAAATAAGAGACAAAGAAATTTGGCCCTGATG	10823						
Db	61	CTATTGGAAGGCTCAACTACGCACTCTGAAACAAAGAAATTTGGCCCTGATG	120						
Qy	1083	GAAATGTTGGTCGTGTGTGTGATTTGAATTGATATCAACAAAGATTTGGTGTGCACTCTTA	11442						
Db	121	GAAATGTTGGTCGTGTGTGTGATTTGAATTGATATCAACAAAGATTTGGTGTGCACTCTTA	180						
Qy	1143	CGTGCAACCAATTTTCCATTAAGACATGTGTTACCCATGAGCTGTTAAACACAAAGCTT	12004						
Db	181	CGTGCAACCAATTTTCCATTAAGACATGTGTTACCCATGAGCTGTTAAACACAAAGCTT	240						
Qy	1203	GGCCCAATGTCGAATATGACATATCTCAAAAGCTTTGGGAATTGAGTGTGTGTGAATG	12624						
Db	241	GGCCCAATGTCGAATATGACATATCTCAAAAGCTTTGGGAATTGAGTGTGTGTGAATG	300						
Qy	1263	GATCAGTGTCTTTCACAAATGTCCTGTATCCATGAAATATCTTAATGTCCTCTCCATG	13222						
Db	301	GATCAGTGTCTTTCACAAATGTCCTGTATCCATGAAATATCTTAATGTCCTCTCCATG	360						
Qy	1323	AAGAGATTAATGCGAGGGAACCCGATATCTGGAATGCTTCAGTCAGAGGAACAGATG	13822						
Db	361	AAGAGATTAATGCGAGGGAACCCGATATCTGGAATGCTTCAGTCAGAGGAACAGATG	420						
Qy	1383	AACGGCTCTGGAGAAACACGTGCACTGCAACAATGAAAGCTACAGCTGTAAACATG	14422						
Db	421	AACGGCTCTGGAGAAACACGTGCACTGCAACAATGAAAGCTACAGCTGTAAACATG	480						
Qy	1443	AAGCAAAATCTGTGCGACGTGGAATGTTATCTCTCATGTGACAAACCACTTGTGAAGAG	15022						
Db	481	AAGCAAAATCTGTGCGACGTGGAATGTTATCTCTCATGTGACAAACCACTTGTGAAGAG	540						
Qy	1503	ACGAAATCTTAATGAGACGTGCTGTGGAGAAATTAATCTTAAATCTGTGTAAAT	15622						
Db	541	ACGAAATCTTAATGAGACGTGCTGTGGAGAAATTAATCTTAAATCTGTGTAAAT	600						
Qy	1563	AGAAATCTGACCACTTAGTAATTAACAATGCTCCATGAGGCGCTTGTCTATTAAT	16222						
Db	601	AGAAATCTGACCACTTAGTAATTAACAATGCTCCATGAGGCGCTTGTCTATTAAT	660						
Qy	1623	AATTTGATTAATTTAATAAATTAAGATGATCTGAAGAGCTGTGATGATATTAAT	16822						
Db	661	AATTTGATTAATTTAATAAATTAAGATGATCTGAAGAGCTGTGATGATATTAAT	720						
Qy	1683	GCTATAGTTAAATGCTTAAAAATTAATTAACCTGTTAATCTTTTTCACCAATCATTT	17442						
Db	721	GCTATAGTTAATGCTTAAAAATTAATTAACCTGTTAATCTTTTTCACCAATCATTT	779						
Qy	1743	----ATTAATTTTTCATTAAGGCAAGTTTCTCTC	1772						
Db	789	TAATAATTTTTCATTAAGGCAAGTTTCTCTC	813						

RESULT 7
 BC012931
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 TITLE
 AUTHORS
 JOURNAL
 COMMENT
 BC012931 1806 bp mRNA linear HTC 04-MAR-2003
 Homo sapiens, Similar to hypothetical protein FLJ73516, clone
 IMAGE:445603, mRNA.
 BC012931
 BC012931.1 GI:15277870
 HTC.
 Homo sapiens (human)
 Homo sapiens
 Bkayapota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Butheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1806)
 Strausberg, R.
 Direct Submission
 Submitted (20-MAR-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk

CDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Guinardene, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louisseged, H.,
 Kowis, C.R., Sneed, A.V., Martin, R.G., Muzny, D.M., Nanavati,
 A.N., Gibbs, R.A.

Clone distribution: MCC clone distribution information can be found through the I.M.A.G.E. Consortium/LMW at: <http://image.llnl.gov>
 Series: IRAX Plate: 28 Row: b Column: 9
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21362047
 This clone has the following problem: retained intron.

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4456033"
/tissue_type="Kidney, hypernephroma"
/clone_lib="NH MGC_89"
/lab_host="DH10B"
/notes=vector:pcmv-sport6"

```

	Query Match	38.4%	Score 680.4	DB 11	Length 1806
	Best Local Similarity	98.6%	Pred. 6.9e-131		
	Matches 707	Conservative	0	Mismatches 7	Indels 3
				Gaps 2	
QY	1058	GGAGACAGAGAAATGCGCCCTGATGAGAGATGCTGTGAGCATGGAATTGTAATAA	111		
Db	1	GGAGACAGAGAAATGCGCCCTGATGAGAGATGCTGTGAGCATGGAATTGTAATAA	60		
QY	1118	CCAAATGATTGGTAGAGATCTTAACTGTGACACCAATATTTCCATTAAGCATGTGTGAC	117		
Db	61	CCAAATGATTGGTAGAGATCTTAACTGTGACACCAATATTTCCATTAAGCATGTGTGAC	120		
QY	1178	CCATGCTGTATTAACACAGACTTGGCCCATGTGCAATATGCACTACCAAGCTTTG	123		
Db	121	CCATGCTGTATTAACACAGACTTGGCCCATGTGCAATATGCACTACCAAGCTTTG	180		
QY	1238	GAATTAGCTGATGATGTGAATGATCACTGTCTTTACAGTCCCTGTATCCAAATGAA	129		

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5303339"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/clone_1b="MYH_MGC_96"
/name="Oryani brain: Vector: plasmidRip (modified)

```

Query Match	38.0%;	Score 673.8;	DB 12;	Length 811;
Best Local Similarity	94.2%;	Pred No. 1.6e-129;		
Matches 745; Conservative	0;	Mismatches 37;	Indels 9;	Gaps 4.

Qy	74	CCGAGGACCTGACATCTGCGGGGAA	CCCTGTGTGCTAGAGG	AGGTGACCTCTCGGCTCCGAGG	133
Ds	22	CCGAGGACCTGACATCTGCGGGGAA	ACCTGTGTGCTAGAGG	AGGTGACCTCTCGGCTCCGAGG	81
Qy	134	TAAGTCGAGAGTCCCGACGTCATCTCAT	CCATCTCTCCCGACCTGAGCAGC	ACTGCTGAG	193
Ds	82	TAAGTCGAGAGTCCCGACGTCATCTCAT	CCATCTCTCTCCCGACCTGAGCAGC	ACTGCTGAG	141
Qy	194	ACCGAGGATCTCTGACGAGGAGGAGG	AGGAGGAGGAGGAGGAGGAGG	AGGAGGAGGAGGAGGAGGAGG	253
Ds	142	ACCA-GGTCTCTGACGAGGAGGAGGAGG	AGGAGGAGGAGGAGGAGGAGGAGG	AGGAGGAGGAGGAGGAGGAGG	200
Qy	254	GCGCGAGCAGAGGAGGAGGAGGAGGAGG	AGGAGGAGGAGGAGGAGGAGGAGG	AGGAGGAGGAGGAGGAGGAGG	313
Ds	201	GCGCGAGCAGAGGAGGAGGAGGAGGAGG	AGGAGGAGGAGGAGGAGGAGGAGG	AGGAGGAGGAGGAGGAGGAGG	260
Qy	314	TCCGAGTCTCTGAGATGATGATCTCT	CTCTGATGATGATGATGATGATGAT	CCGAGGAGGAGGAGGAGGAGG	373
Ds	261	TCCGAGTCTCTGAGATGATGATCTCT	CTCTGATGATGATGATGATGATGAT	CCGAGGAGGAGGAGGAGGAGG	320
Qy	374	GGGGCTGAGAGAGGTGTGAGACCGCGTA	CCCTCAAGTGTCTGTGCGGGATT	CCGAGAGCGGAA	433
Ds	321	GGGGCTGAGAGAGGTGTGAGACCGCGTA	CCCTCAAGTGTCTGTGCGGGATT	CCGAGAGCGGAA	380
Qy	434	GTCGACCCGATAGGTGTGTGAGAGGAG	AGGAGGAGGAGGAGGAGGAGGAGG	AGGAGGAGGAGGAGGAGGAGG	493
Ds	381	GTCGACCCGATAGGTGTGTGAGAGGAG	AGGAGGAGGAGGAGGAGGAGGAGG	AGGAGGAGGAGGAGGAGGAGG	440
Qy	494	GAGCGTGTGAGTGTGAGTCTGTGATCCG	CCGAGGAGGAGGAGGAGGAGGAGG	AGGAGGAGGAGGAGGAGGAGG	553
Ds	441	GAGCGTGTGAGTGTGAGTCTGTGATCCG	CCGAGGAGGAGGAGGAGGAGGAGG	AGGAGGAGGAGGAGGAGGAGG	500
Qy	554	CCGAGACGCAATTTCAAGGTGTGAGGAG	CGGTTTGTGAGGAGGAGGAGGAGG	AGGAGGAGGAGGAGGAGGAGG	613
Ds	501	CCGAGACGCAATTTCAAGGTGTGAGGAG	CGGTTTGTGAGGAGGAGGAGGAGG	AGGAGGAGGAGGAGGAGGAGG	560
Qy	614	GCCCTCATCCAAACCGGCGGAGGAGG	CTGACCTTGTGACAGAGATCATCTG	AGCTTATGAG	673
Ds	561	GCCCTCATCCAAACCGGCGGAGGAGG	CTGACCTTGTGACAGAGATCATCTG	AGCTTATGAG	620
Qy	674	AGATGGGCGTCTGAGAGCGCTCATCTTAA	---CTTCCCGGAGACCGGCATGAG---	GTC	727
Ds	621	AGATGGGCGTCTGAGAGCGCTCATCTTAA	---CTTCCCGGAGACCGGCATGAG---	GTC	680
Qy	728	ATCCCATCTCTCAACCGGAGTCACTAGAC	CAATGTGTGATC---ATGATCGGACATCTGA		785
Ds	681	ATCCCATCTCTCAACCGGAGTCACTAGAC	CAATGTGTGATC---ATGATCGGACATCTGA		740
Qy	786	AAGGCAAAAATTTCTGCATCTATTCAAAG	AGGCTACAGGTGACATAGGCTATAGAG		845
Ds	741	AAGGCAAAAATTTCTGCATCTATTCAAAG	AGGCTACAGGTGACATAGGCTATAGAG		800
Qy	846	TAGGAGAAAAA	856		
Ds	801	GTAGGGGAAA	811		

BI600419 1000 bp mRNA linear EST 07-SEP-2001
 LOCUS 603246458P1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5288774 5',
 DEFINITION mRNA sequence.
 ACCESSION BI600419
 VERSION BI600419.1 GI:15493358
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (Bases 1 to 1000)
 NIH-MGC http://mgs.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLES Unpublished
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cga@dbp-remail.nih.gov
 Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: L1M11730 row: c column: 15
 High quality sequence stop: 701.
 Location/Qualifiers
 1..1000
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5288774"
 /tissue_type="hypothalamus"
 /lab_host="DB108"
 /clone_lib="NIH_MGC_96"
 /note="Organ: Brain; Vector: pBluescript (modified
 pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcagag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
 size-selected for average insert size 2.3 kb and
 normalized to ROT 5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NIH/NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."
 BASE COUNT 206 a 299 c 321 g 174 t
 ORIGIN
 Query Match 36.9%; Score 654.2; DB 12; Length 1000;
 Best Local Similarity 93.7%; Pired. No. 1.9e-125;
 Matches 747; Conservative 0; Mismatches 43; Indels 7; Gaps 6;
 QY 74 CCGAGAGAGCTGATCTGCGCAACCTGTGTGTGACGTCACGCTCTGCTGCTCCAG 133
 DB 22 CCGAGAGAGCTGATCTGCGCAACCTGTGTGTGACGTCACGCTCTGCTGCTCCAG 81
 QY 134 TAGCTGCAAGCTCCCGACGTCACCTGCTCCCGACCTGCGCGACCTGCTGAG 123
 DB 82 TAGCTGCAAGCTCCCGACGTCACCTGCTCCCGACCTGCGCGACCTGCTGAG 141
 QY 194 ACCAGAGTCTGCAAGGCTGAGAGGCGCTGCAAGGAGGCTGAGAGGACTGAGAGC 253
 DB 142 ACCA-GGTCTGCAAGGCTGAGAGGCGCTGCAAGGAGGCTGAGAGGACTGAGAGC 199
 QY 254 GCGCGGCGCTGAGAGGCGCTGAGAGGCGCTGAGAGGCGCTGAGAGGCGCTGAGAGC 313
 DB 200 GCGCGGCGCTGAGAGGCGCTGAGAGGCGCTGAGAGGCGCTGAGAGGCGCTGAGAGC 259
 QY 314 TCCAGATTGCTGAGAGGCTGAGAGGCGCTGAGAGGCGCTGAGAGGCGCTGAGAGC 373
 DB 260 TCCAGATTGCTGAGAGGCTGAGAGGCGCTGAGAGGCGCTGAGAGGCGCTGAGAGC 319
 QY 374 GGGGCTGAAAGAGTGTGAGCGCGTACTCAAGTGTCTGAGCGGGTTCCGCAACGAGGA 433

DB 320 GGGGCTGAAAGAGTGTGAGCGCGTACTCAAGTGTCTGAGCGGGTTCCGCAACGAGGA 379
 QY 434 GTGAACCGTACGATGTGAGAGGCTGAGAGGCGCTGAGAGGCGCTGAGAGGCGCTGAGAGC 493
 DB 380 GTGAACCGTACGATGTGAGAGGCTGAGAGGCGCTGAGAGGCGCTGAGAGGCGCTGAGAGC 439
 QY 494 GAGCTGTGAGTGTGAGAGGCTGAGAGGCGCTGAGAGGCGCTGAGAGGCGCTGAGAGC 553
 DB 440 GAGCTGTGAGTGTGAGAGGCTGAGAGGCGCTGAGAGGCGCTGAGAGGCGCTGAGAGC 499
 QY 554 CCGCAACGAAATTCACGATGCGCCACGATGCGAGAGGCGCTGAGAGGCGCTGAGAGC 612
 DB 500 CCGCAACGAAATTCACGATGCGCCACGATGCGAGAGGCGCTGAGAGGCGCTGAGAGC 559
 QY 613 GGCCTCTATCCACGATGCGCCACGATGCGAGAGGCGCTGAGAGGCGCTGAGAGC 671
 DB 560 GGCCTCTATCCACGATGCGCCACGATGCGAGAGGCGCTGAGAGGCGCTGAGAGC 619
 QY 672 AGAGATGGGCGCTGAGAGGCGCTGAGAGGCGCTGAGAGGCGCTGAGAGGCGCTGAGAGC 731
 DB 620 AGAGATGGGCGCTGAGAGGCGCTGAGAGGCGCTGAGAGGCGCTGAGAGGCGCTGAGAGC 678
 QY 732 CAGTGTCTGACCGGCTGAGAGGCGCTGAGAGGCGCTGAGAGGCGCTGAGAGGCGCTGAGAGC 791
 DB 679 CAGTGTCTGACCGGCTGAGAGGCGCTGAGAGGCGCTGAGAGGCGCTGAGAGGCGCTGAGAGC 736
 QY 792 CAGTGTCTGACCGGCTGAGAGGCGCTGAGAGGCGCTGAGAGGCGCTGAGAGGCGCTGAGAGC 851
 DB 739 CAGTGTCTGACCGGCTGAGAGGCGCTGAGAGGCGCTGAGAGGCGCTGAGAGGCGCTGAGAGC 796
 QY 852 AAAAATGCTGATCTGAGAGGCGCTGAGAGGCGCTGAGAGGCGCTGAGAGGCGCTGAGAGC 868
 DB 797 AAAAATGCTGATCTGAGAGGCGCTGAGAGGCGCTGAGAGGCGCTGAGAGGCGCTGAGAGC 813
 RESULT 10
 AV690200 676 bp mRNA linear EST 16-JAN-2002
 LOCUS AV690200 GRC Homo sapiens cDNA clone GRCAX02 5', mRNA sequence.
 DEFINITION AV690200 GRC Homo sapiens cDNA clone GRCAX02 5', mRNA sequence.
 ACCESSION AV690200
 VERSION AV690200.1 GI:10292063
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (Bases 1 to 676)
 Xu, X., Huang, J., Xu, Z., Qian, B., Zhu, Z., Yan, Q., Cai, T., Zhang, X.,
 Xiao, H., Qu, J., Liu, F., Huang, Q., Cheng, Z., Li, N., Du, J., Hu, W.,
 Shen, K., Lu, G., Fu, G., Zhong, M., Xu, S., Gu, W., Huang, W., Zhao, X.,
 Hu, G., Gu, J., Chen, Z., and Han, Z.
 Insight into hepatocellular carcinogenesis at transcriptome level
 by comparing gene expression profiles of hepatocellular carcinoma
 with those of corresponding noncancerous liver
 Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
 21625106
 MEDLINE
 PUMED 11752456
 COMMENT Contact: Zenguan Han
 Chinese National Human Genome Center at Shanghai
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
 201203, P. R. China
 Tel: 86-21-50801919 (ex.45)
 Fax: 86-21-50801922
 Email: hanzg@chgc.sh.cn
 This clone is available at CHGC in Shanghai.
 Location/Qualifiers
 1..676
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="GRCAX02"
 /tissue_type="hepatocellular carcinoma"


```

/dev_stage="Adult"
/lab_host="SOLR"
/clone_lib="NGC"
/notes="Vector: pbluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

```

```

BASE COUNT      94 a 223 c 231 g 125 t      3 others
ORIGIN
Query Match      36.3%; Score 644.2; DB 9; Length 676;
Best Local Similarity 98.1%; Pred. No. 2.2e-123;
Matches 661; Conservative 0; Mismatches 11; Indels 2; Gaps 1;

```

```

QY 74 CCGAGAGCTGCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 133
Db 4 CCGAGAGCTGCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 63
QY 134 TAGCTCGAGCTCCCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 193
Db 64 TAGCTCGAGCTCCCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 123
QY 194 ACCAGAGCTCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 253
Db 124 ACCAGAGCTCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 183
QY 254 GCGCGCGCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 313
Db 184 GCGCGCGCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 243
QY 314 TCCAGATGCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 373
Db 244 TCCAGATGCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 303
QY 374 GCGCGCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 433
Db 304 GCGCGCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 363
QY 434 GTGAGAGCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 493
Db 364 GTGAGAGCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 423
QY 494 GAGCTGCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 553
Db 424 GAGCTGCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 483
QY 554 CCGAGAGCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 613
Db 484 CCGAGAGCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 543
QY 614 GCGCTGCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 673
Db 544 GCGCTGCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 603
QY 674 AGATGAGCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 733
Db 604 AGATGAGCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 661
QY 734 ATGCTGCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 747
Db 662 ATGCTGCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 675

```

```

RESULT 11
BI601096      773 bp      mRNA      linear      EST 07-SEP-2001
LOCUS      603249588B1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5301380 5'
DEFINITION      mRNA sequence.
ACCESSION      BI601096
VERSION      BI601096.1 GI:15494035
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

```

```

REFERENCE      1 (bases 1 to 773)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished
COMMENT      Contact: Robert Strausberg, Ph.D.
            Email: cga@bbs-remail.nih.gov
            Tissue Procurement: Miklos Balke, M.D., Ph.D.
            CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/BLML at:
            http://image.lnl.gov
            Plate: LMML1762 row: p column: 21
            High quality sequence stop: 759.
            Location/Qualifiers

```

```

FEATURES
source
1..773
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5301380"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_96"
/notes="Organ: brain; Vector: pbluescript (modified
pbluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (GTCGAG
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTN-3',
size-selected for average insert size 2.3 kb and
normalized to 10^5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation) Library
constructed by M. Brownstein (NHGRI/NHGLI, National
Institutes of Health). Note: this is a NIH-MGC Library."

```

```

BASE COUNT      130 a 242 c 254 g 146 t      1 others
ORIGIN
Query Match      36.1%; Score 641; DB 12; Length 773;
Best Local Similarity 96.3%; Pred. No. 1e-122;
Matches 720; Conservative 0; Mismatches 21; Indels 7; Gaps 6;

```

```

QY 74 CCGAGAGCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 133
Db 22 CCGAGAGCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 81
QY 134 TAGCTGCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 193
Db 82 TAGCTGCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 141
QY 194 ACCAGAGCTCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 253
Db 142 ACCAGAGCTCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 200
QY 254 GCGCGCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 313
Db 201 GCGCGCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 260
QY 314 TCCAGATGCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 373
Db 261 TCCAGATGCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 320
QY 374 GCGCGCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 433
Db 321 GCGCGCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 380
QY 434 GTGAGAGCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 493
Db 381 GTGAGAGCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 440
QY 494 GAGCTGCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 553
Db 441 GAGCTGCTGCTGCGCAACTGTGTCTGACGCTACGCTCTGCTGCGAGC 500

```

QY	554	CCGACACAGAAATTTAC-GGTGGCCACGGTTTGGGGAAGC-ACCCTGCAGCTCTTGG-	610
DB	501	CCGACACAGAAATTTACAGGGTGGCCACGGANTTGGGAGAGAACCGCTGCAGCTCTTGG-	560
QY	611	TTGGCCCTTCATCCACCGCGGGGGCTGCACCTTGGCAGACAAAGATTCATCTGCTTAT	670
DB	561	TTGGCCCTTCATCCACCGCGGGGGCTGCACCTTGGCAGACAAAGATTCATCTGCTTAT	620
QY	671	GAGAGATGGGGCTTGGAGCCGCTCATCTTTAACTTCCCGGAGACC--GCATGAGGTCA	728
DB	621	GAGAGAGGGGGCTTGGAGCCGCTCATCTTTAACTTCCCGGAGACC--GCATGAGGTCA	680
QY	729	TCCCATGCTTCACCGGGGTGCAGTACATTT-GTTCATATGATTCGCAATCTGAAA	787
DB	681	TCCCATGCTTCACCGGGGTGCAGTACATTTGATTCATGATTCGCAATCTGCA	740
QY	788	GGCAGAAAATTCGCAATCTTATTTCAA	815
DB	741	AGGCCAAAATTCGCAATCTTATTTCAA	768
RESULT 12			
AK008312			
LOCUS	840 bp	mRNA	linear HTC 05-DEC-2002
DEFINITION	Mus musculus adult male small intestine cDNA, RIKEN full-length enriched library, clone:2010104G21 product:GRAT1 homolog [Homo sapiens], full insert sequence.		
ACCESSION	AK008312		
VERSION	AK008312.1	GI:12842420	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Mammalia; Eutheria; Rodentia; Scuriognath; Muridae; Murinae; Mus. Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scuriognath; Muridae; Murinae; Mus.		
AUTHORS	Carninci, P. and Hayashizaki, Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		
REFERENCE	2		
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
MEDLINE	20493374		
PUBMED	11042159		
REFERENCE	3		
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komno, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, K., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kiz, A., and Hayashizaki, Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		
MEDLINE	20530913		
PUBMED	11076861		
REFERENCE	4		
AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, T., Hara, A., Fukunishi, Y., Komno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamataka, I., Saito, T., Okazaki, Y., Gotohori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiya, H., Kuehl, P., Lewis, S., Matsuo, Y., Nishida, T., Reese, G., Quackenbush, J., Schriml, L.M., Staudt, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barish, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bull, C.,		
FEATURES	source		
CDS	Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamaly, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzerelli, J., Mombauts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakumoto, N., Sasaki, H., Sato, K., Schomach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyokura, K., Wang, K.H., Weitz, C., Whitaker, C., Wilmink, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohetsuki, S. and Hayashizaki, Y.		
TITLE	Functional annotation of a full-length mouse cDNA collection		
JOURNAL	Nature 409 (6821), 685-690 (2001)		
MEDLINE	21085660		
PUBMED	11217851		
REFERENCE	5		
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.		
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs		
JOURNAL	Nature 420, 563-573 (2002)		
MEDLINE	6 (bases 1 to 840)		
PUBMED	Adachi, J., Aizawa, K., Akabira, S., Akiyama, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Imoto, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Komno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, K., Sakai, C., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, I., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Yamamoto, M. and Hayashizaki, Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsunumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-resgsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)		
COMMENT	Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.		
TITLE	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGACAGAGATCCAGAGCTCTTTTCTTTTCTTAA 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 5.0 and subtraction to Rot = 20.0. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGACAGATCCAGAGCTCTTTTCTTTTCTTAA 3']. cDNA was cleaved with XhoI and SclI. Cloning sites, 5' end: XhoI; 3' end: SclI. Host: SOLR.		
FEATURES	location/qualifiers		
source	1..840		
	/organism="Mus musculus"		
	/mol type="mRNA"		
	/strain="C57BL/6J"		
	/db_xref="FANTOM DB:2010104G21"		
	/db_xref="MGI:1855748"		
	/db_xref="taxon:10090"		
	/clone="2010104G21"		
	/sex="male"		
	/tissue type="small intestine"		
	/clone lib="RIKEN full-length enriched mouse cDNA library"		
	/dev stage="adult"		
	131...>838		
	/note="unnamed protein product, GRAT1 homolog [Homo sapiens] (SPTK106873, evidence: FASTA, 93.2%ID, 100%length, match=1284)		
	putative"		

Query Match	35.1%;	Score 622.4;	DB 10;	Length 906
Best Local, Similarity	96.0%;	Pred. No. 7.6e-119;		

REFERENCE
AUTHORS
TITLE

1 (bases 1 to 683)
Xu, X., Huang, Y., Xu, Z., Qian, B., Zhu, Z., Yan, Q., Cai, T., Zhang, X., Xiao, H., Ou, J., Liu, F., Huang, Q., Cheng, Z., Li, N., Du, J., Hu, W., Shen, K., Lu, G., Fu, G., Zhong, M., Xu, S., Gu, W., Huang, W., Zhao, X., Hu, G., Gu, Y., Chen, Z., and Han, Z.
Insight into hepatocellular carcinogenesis at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma

JOURNAL
MEDLINE
PUBMED
COMMENT

with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
21625106
11752456
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES

source

Location/Qualifiers

1. .683
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GKCALB10"
/tissue_type="hepatocellular carcinoma"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_1ib="GKC"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 98 a 222 c 228 g 135 t
ORIGIN

Query Match

Best Local Similarity 97.0%; Score 622.2; DB 9; Length 683;
Matches 655; Conservative 0; Mismatches 18; Indels 2; Gaps 2;

QY 74 CCGAGAGCTGCACTCTGCGGCAACTGTGTCTGAGCTTACGTCTCTGCTCCGACG 133
DB 4 CCGAGAGCTGCACTCTGCGGCAACTGTGTCTGAGCTTACGTCTCTGCTCCGACG 63
QY 134 TAGCTCGAGCTCTCCGAGCTCTCACTCCATTCCTTCCGACCTGCGCGCGACCTGCTCAAG 193
DB 64 TAGCTCGAGCTCTCCGAGCTCTCACTCCATTCCTTCCGACCTGCGCGCGACCTGCTCAAG 123
QY 194 ACCAGGCTCTGCCAAGCGCTAGAGGCGCGGTGCGACGAGGCGCTTAGGGAATGCGGAGC 253
DB 124 ACCAGGCTCTGCCAAGCGCTAGAGGCGCGGTGCGACGAGGCGCTTAGGGAATGCGGAGC 183
QY 254 GCGCGCGCTATGGGCGCGCGCTCTGGGCGCGGCTCTCTCCGCGGTGGCTGCGGCTTT 313
DB 184 GCGCGCGCTATGGGCGCGCGCTCTGGGCGCGGCTCTCTCCGCGGTGGCTGCGGCTTT 243
QY 314 TCCAGATTGTGCGATGATGCTCTCTGCTGGCTTGAAGTCCGACGACCCGAGTTCCCG 373
DB 244 TCCAGATTGTGCGATGATGCTCTCTGCTGGCTTGAAGTCCGACGACCCGAGTTCCCG 303
QY 374 GGGGCTGAAGCAATGTGAGACCGGTACCTCAAGCTGTCTGCGGCTTCCGACACGGA 433
DB 304 GGGGCTGAAGCAATGTGAGACCGGTACCTCAAGCTGTCTGCGGCTTCCGACACGGA 363
QY 434 GTGAACCGTACGCTGTGAGCTGAGCGAGGAGGCGGTACAGGACGACGCGCTG 493
DB 364 GTGAACCGTACGCTGTGAGCTGAGCGAGGAGGCGGTACAGGACGACGCGCTG 423
QY 494 GAGCTGTGCTGGGCTCTGTGACCGCGACGAGGCGCGGCGCTTAAAGCTGTAC 553
DB 424 GAGCTGTGCTGGGCTCTGTGACCGCGACGAGGCGCGGCGCTTAAAGCTGTAC 483
QY 554 CCGCAACGAATTTCAAGGTGCGCACGCTTGGGGAAGACCGTGCAATCTTGTGTTG 613
DB 484 CCGCAACGAATTTCAAGGTGCGCACGCTTGGGGAAGACCGTGCAATCTTGTGTTG 543
QY 614 GCGCTATTCGAAGCGGCGGCGGCGCACTTCGAGACAAATTCATCTGGCTATGAG 673
DB 544 GCGCTATTCGAAGCGGCGGCGGCGCACTTCGAGACAAATTCATCTGGCTATGAG 603
QY 674 -AGATGGCGCTGTGAGCGGTCACTTTAACTTCCCGGAGACCGGCAATGAGTCAATCC 732
DB 603 -AGATGGCGCTGTGAGCGGTCACTTTAACTTCCCGGAGACCGGCAATGAGTCAATCC 702

DB 604 AAGAAGGAGGCTGTGAGCGCGGCACTTTACTTCCCGAG-CGCCAATGAGGTCAATCC 662
QY 733 CATGCTCAACCGG 747
DB 663 CATGTTTACCGG 677

Search completed: October 6, 2003, 17:37:45
Job time : 3893 secs


```

Db      1441 TGAAGCAAAATCTGGGAGGAGTGAATTTCTCATGTGAGCAACCAACCTTGAAGA 1500
Qy      1501 AGAGCAAACTCTTAATCAAGAGCTGCTGTGCAAAATTAATCTTAATCTGTGTA 1560
Db      1501 AGAGCAAACTCTTAATCAAGAGCTGCTGTGCAAAATTAATCTTAATCTGTGTA 1560
Qy      1561 ATAGAAACCTGAGCAACATTAAGAAATTAACAGACTGCAATGAGGACCTAATTTCTATTA 1620
Db      1561 ATAGAAACCTGAGCAACATTAAGAAATTAACAGACTGCAATGAGGACCTAATTTCTATTA 1620
Qy      1621 TAAATGATTAATTTAATTAATAATTAAGAGTGAATCTGAAGTCTGAGATGATATAT 1680
Db      1621 TAAATGATTAATTTAATAATAATTAAGAGTGAATCTGAAGTCTGAGATGATATAT 1680
Qy      1681 ATGCTATGATTAATAAGGCTTAAATTAATTAACCTGTAACCTTTTCCACCAACTCA 1740
Db      1681 ATGCTATGATTAATAAGGCTTAAATTAATTAACCTGTAACCTTTTCCACCAACTCA 1740
Qy      1741 TTATATATATTTTTCATAGGCAAGTTCTCTCAG 1774
Db      1741 TTATATATATTTTTCATAGGCAAGTTCTCTCAG 1774

RESULT 2
AA15995
ID      AA15995 standard; cDNA; 2773 BP.
XX
AC      AA15995;
XX
DT      12-JUN-2000 (first entry)
XX
DE      Human protein clone HP10574 full length coding sequence.
XX
KW      Human protein; hydrophobic domain; nutritional source; haematopoiesis;
KW      cytokine production; cell proliferation; cell differentiation;
KW      immune deficiency; infectious disease; autoimmune disorder; asthma;
KW      multiple sclerosis; systemic lupus erythematosus; rheumatoid arthritis;
KW      allergic reaction; osteoporosis; osteoarthritis; periodontal disease;
KW      nervous system disorder; Alzheimer's disease; Parkinson's disease;
KW      Huntington's disease; liver fibrosis; lung fibrosis; reperfusion injury;
KW      systemic cytokine damage; tissue differentiation; contraceptive; stroke;
KW      coagulation disorder; myocardial infarction; inflammatory condition;
KW      septic shock; sepsis; ischaemia; reperfusion injury; arthritis; tumour;
KW      nephritis; therapy; ss.
XX
OS      Homo sapiens.
XX
PN      WO200005367-A2.
XX
PD      03-FEB-2000.
XX
PF      22-JUL-1999; 99WO-JP03929.
XX
PR      24-JUL-1998; 98JP-0208820.
PR      07-AUG-1998; 98JP-0224105.
PR      25-AUG-1998; 98JP-0238116.
PR      09-SEP-1998; 98JP-0254736.
PR      23-SEP-1998; 98JP-0275505.
XX
PA      (SAGA) SAGAMI CHEM RES CENT.
PA      (PROT-) PROTEGENE INC.
XX
PI      Kato S, Kimura T;
XX
DR      MPI; 2000-182694/16.
DR      P-PSDB; AAY94897.
XX
PT      Novel human proteins having hydrophobic domains useful for treating
PT      osteoporosis, Alzheimer's disease, Parkinson's disease, asthma,
PT      multiple sclerosis, rheumatoid arthritis, cancer, anaemia, and stroke -
XX
PS      Claim 4; Page 348-351; 351pp; English.

```

```

XX      This sequence encodes a human protein of the invention, which has
CC      hydrophobic domains. The DNA sequences can be used as a probe or as a
CC      genetic marker. The protein can also be used as a marker, and to identify
CC      potential genetic disorders. The DNA and protein can also be used as
CC      nutritional sources or supplements. The protein exhibits cytokine, cell
CC      proliferation, cell differentiation activities and induces production of
CC      other cytokines in certain cell populations. The protein also exhibits
CC      immune stimulating or immune suppressing activity. It can be used in the
CC      treatment of various immune deficiencies and disorders, and to treat
CC      infectious diseases caused by viral, bacterial, fungal or other
CC      infections. The protein is also used for treating autoimmune disorders
CC      such as multiple sclerosis, systemic lupus erythematosus, and rheumatoid
CC      arthritis. It is also useful in the treatment of allergic reactions and
CC      conditions such as asthma, and in immune suppression after organ
CC      transplantation. The protein is useful in regulation of haematopoiesis
CC      and consequently in the treatment of myeloid or lymphoid cell or
CC      deficiencies. It is also used in compositions for tissue growth or
CC      regeneration. The protein is also used in the treatment of osteoporosis
CC      or osteoarthritis and in the treatment of periodontal disease and other
CC      tooth repair processes. The protein is used in the treatment of nervous
CC      system disorders such as Alzheimer's disease, Parkinson's disease, and
CC      Huntington's disease. They are useful for protection or regeneration and
CC      treatment of lung or liver fibrosis, reperfusion injury in various
CC      tissues, and conditions resulting from systemic cytokine damage. They are
CC      also used for promoting or inhibiting tissue differentiation. They are
CC      also used as contraceptives since they exhibit activin or inhibin related
CC      activities and as a fertility inducing therapeutic. They are used for
CC      treating various coagulation disorders and in treatment and prevention of
CC      conditions resulting from coagulation activities e.g. myocardial
CC      infarction or stroke. They also acts as receptors, receptor ligands or
CC      inhibitors or agonists of receptor/ligand interactions. They are used to
CC      treat inflammatory conditions such as septic shock, sepsis, ischaemia
CC      reperfusion injury, arthritis, and nephritis. They can be used to
CC      prevent tumours.
XX
SQ      Sequence 2773 BP; 803 A; 545 C; 617 G; 808 T; 0 other;
XX
Query Match      93.7%; Score 1662.8; DB 21; Length 2773;
Best Local Similarity      99.4%; Pred. No. 0;
Matches 1690; Conservative 0; Mismatches 8; Indels 3; Gaps 2;
Qy      74 CCGAGAGACTGCATCTGCGGCAACCTGTGTGCTGACGCTACGCTCTGCTGCTCCAG 133
Db      22 CCGAGAGACTGCATCTGCGGCAACCTGTGTGCTGACGCTACGCTCTGCTGCTCCAG 81
Qy      134 TAGCTGCACTGCCAGTCCAGTCTCACTCCATTCCTTCCCACTTGGCGGCACTGCTCAAG 193
Db      82 TAGCTGCACTGCCAGTCCAGTCTCACTCCATTCCTTCCCACTTGGCGGCACTGCTCAAG 141
Qy      194 ACCAGAGTCCGCAAGCGCTAGAGAGGCGGCTGACGAGGAGCTAGAGGAACTGAGGAGC 253
Db      142 ACCAGAGTCCGCAAGCGCTAGAGAGGCGGCTGACGAGGAGCTAGAGGAACTGAGGAGC 201
Qy      254 GCGGCGCATGAGGAGCGCGGCTGAGGAGCGGAGGCTCTCTGCGCGGTGCTGCGGCTTT 313
Db      202 GCGGCGCATGAGGAGCGCGGCTGAGGAGCGGAGGCTCTCTGCGCGGTGCTGCGGCTTT 261
Qy      314 TCGAATTTGCTGCAATGCTGCTTCTGCTGCGGCTGAGTCCGAGGCACTCCGATTCGCGG 373
Db      262 TCGAATTTGCTGCAATGCTGCTTCTGCTGCGGCTGAGTCCGAGGCACTCCGATTCGCGG 321
Qy      374 GGGGCTGAAGAGTGTGACCGGCTTACCTCAACGCTGCTGAGGAGTCTCCGACACGAGGA 433
Db      322 GGGGCTGAAGAGTGTGACCGGCTTACCTCAACGCTGCTGAGGAGTCTCCGACACGAGGA 381
Qy      434 GTGACCTGTAAGTGTGTGGAGTGTGAGCGAGAGAGGCGTGTACCGGCAAGACTGCGCGTG 493
Db      382 GTGACCTGTAAGTGTGTGGAGTGTGAGCGAGAGAGGCGTGTACCGGCAAGACTGCGCGTG 441
Qy      494 GAGCCTGTGAGTGGGCTCTGTGTAACGCGCAACGAGGAGCGCGGAGCTTAAACGCTGTAAAC 553
Db      442 GAGCCTGTGAGTGGGCTCTGTGTAACGCGCAACGAGGAGCGCGGAGCTTAAACGCTGTAAAC 501

```

```

QY 554 CCGGACACGAAATTTCAGGTCGCCAGGTTTGGGGAAGCAACGACAGCTCTTGGTTG 613
DB 502 CCGGACACGAAATTTCAGGTCGCCAGGTTTGGGGAAGCAACGACAGCTCTTGGTTG 561
QY 614 GCCCTCATCCAAACGACGCGGCGGCGCTGCACTTGGACAGCAAGATCATCTGGCTTATG 673
DB 562 GCCCTCATCCAAACGACGCGGCGGCGCTGCACTTGGACAGCAAGATCATCTGGCTTATG 621
QY 674 AGATGGGCGCTTGGACCGCTCATCTTTAATCTCCCGGGAACCCGATAGAGTCATCCG 733
DB 622 AGAGGGCGCTTGGACCGCTCATCTTTAATCTCCCGGGAACCCGATAGAGTCATCCG 681
QY 734 ATGTCTACCCGCGGTGCAATGACATTTGTTCATCATATGCGCAATCGAAGGACA 793
DB 682 ATGTCTACCCGCGGTGCAATGACATTTGTTCATCATATGCGCAATCGAAGGACA 741
QY 794 AAAATTCGCAATCTATTCGAAAGGCAACAAGTGAAGATGCAAGTAAAGGAAA 853
DB 742 AAAATTCGCAATCTATTCGAAAGGCAACAAGTGAAGATGCAAGTAAAGGAAA 801
QY 854 AAAATTCGCAATCTATTCGAAAGGCAACAAGTGAAGATGCAAGTAAAGGAAA 913
DB 802 AAAATTCGCAATCTATTCGAAAGGCAACAAGTGAAGATGCAAGTAAAGGAAA 861
QY 914 ATTAAGGCGGCAACGTGGCTATTTATCTTTTATCTGTCGAAAGCTACGGAATGCA 973
DB 862 ATTAAGGCGGCAACGTGGCTATTTATCTTTTATCTGTCGAAAGCTACGGAATGCA 921
QY 974 AGAGCTCAAGCGAGAGAGAGAGGCAATTAAGGCAAGTCAAAAAGCTATGGAAG 1033
DB 922 AGAGCTCAAGCGAGAGAGAGAGGCAATTAAGGCAAGTCAAAAAGCTATGGAAG 981
QY 1034 CTTCACACTACGCACTGTGAAACAGAGAGACAGGAAATGGCCCTGATGAGATGTTGT 1093
DB 982 CTTCACACTACGCACTGTGAAACAGAGAGACAGGAAATGGCCCTGATGAGATGTTGT 1041
QY 1094 GCTGTGTCATTTGATTTGATTAACCAATGATTTGGTACGATCTTAAGTGAACCT 1153
DB 1042 GCTGTGTCATTTGATTTGATTAACCAATGATTTGGTACGATCTTAAGTGAACCT 1101
QY 1154 ATTTCCATAAGCATGTGTGACCCATGCTGTTAAACAAGACTTGGCCCATGTGC 1213
DB 1102 ATTTCCATAAGCATGTGTGACCCATGCTGTTAAACAAGACTTGGCCCATGTGC 1161
QY 1214 AAATGTGACATCTCAAGCTTGGGAATGAGTGAAGTGAAGATGATGATGCTCT 1273
DB 1162 AAATGTGACATCTCAAGCTTGGGAATGAGTGAAGTGAAGATGATGATGCTCT 1221
QY 1274 TTACAAGTCCCTGATTCACATGAATATCTAATAGTGGCCCTCCCATGAAGAGATAT 1333
DB 1222 TTACAAGTCCCTGATTCACATGAATATCTAATAGTGGCCCTCCCATGAAGAGATAT 1281
QY 1334 CCGACCGAGACCGGATCATCTGATATCTTCAATGACAGGAAACAGATGAACCGCTCTG 1393
DB 1282 CCGACCGAGACCGGATCATCTGATATCTTCAATGACAGGAAACAGATGAACCGCTCTG 1341
QY 1394 GAGGAAACCGTGAAGTCAACAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1453
DB 1342 GAGGAAACCGTGAAGTCAACAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1401
QY 1454 GTGGCAGTGAATTTATCTCATGTGACACCAACCTTTGAAGAAAGCAAACTCTCT 1513
DB 1402 GTGGCAGTGAATTTATCTCATGTGACACCAACCTTTGAAGAAAGCAAACTCTCT 1461
QY 1514 AATCAAGAGACTGTGTGAGAAATTAATCTTAATCTGTGAAATTAAGAACTTGA 1573
DB 1462 AATCAAGAGACTGTGTGAGAAATTAATCTTAATCTGTGAAATTAAGAACTTGA 1521
QY 1574 ACCATTAGTAATTAACAAGACTGCAATCGAGGCTAGTTTCTATTAATTAATGATTA 1633
DB 1522 ACCATTAGTAATTAACAAGACTGCAATCGAGGCTAGTTTCTATTAATTAATGATTA 1581

```

```

QY 1634 TTATAATAATTAAGAGTGAATCTGAAGAGTCCAGATGACTAATTAATGCTAATTA 1693
DB 1582 TTATAATAATTAAGAGTGAATCTGAAGAGTCCAGATGACTAATTAATGCTAATTA 1641
QY 1694 AATGGCTTAATAATTAATTAACCTGTTAATCTTTTTCACCAAACTCATTAATATTTT 1753
DB 1642 A--TGGCTTAATAATTAATTAACCTGTTAATCTTTTTCACCAAACTCATTAATATTTT 1698
QY 1754 CATAGGCAAGTTCCTCTCAG 1774
DB 1698 CATAGGCAAGTTCCTCTCAG 1719

```

RESULT 3

AAI72321 standard; cDNA; 2796 BP.

AAI72321; 15-APR-2002 (first entry)

ISIGP-4 cDNA.

Human; intracellular signalling protein; ISIGP; gene; cell proliferation; autoimmune; inflammation; gastrointestinal disorder; reproductive disorder; developmental disorder; ss.

Homo sapiens.

Key Location/Qualifiers
 CDS 217..1503
 FT /*tag= a
 FT /product= "ISIGP-4"

W0200194391-A2.

13-DEC-2001.

07-JUN-2001; 2001WO-US18595.

08-JUN-2000; 2000US-210582P.

16-JUN-2000; 2000US-212443P.

(IMCY-) INCYTE GENOMICS INC.

Yue H, He A, Nguyen DB, Yao MG, Bandman O, Burford N, Tang YT; Ku Y, Hafalia A, Azimzai Y, Walia NK;

WPI; 2002-154564/20.
 P-PSDB; AAB47874.

New human intracellular signalling protein and polynucleotides useful for diagnosing, treating or preventing cell proliferative, autoimmune/inflammatory, gastrointestinal, reproductive and developmental disorders -

Claim 5; Page 104; 106pp; English.

The sequences given in AAI72318-22 encode novel human intracellular signalling proteins (ISIGP). The polynucleotides and ISIGP proteins may be used for the diagnosis, treatment or prevention of cell proliferative, autoimmune/inflammatory, gastrointestinal, reproductive and developmental disorders. The protein encoded by this sequence has homology to mouse gl-related zinc finger protein.

Sequence 2796 BP; 821 A; 548 C; 620 G; 807 T; 0 other;

Query Match 93.7%; Score 1662.8; DB 24; Length 2796;
 Best Local Similarity 99.4%; Pred. No. 0;

Matches 1690; Conservative 0; Mismatches 8; Indels 3; Gaps 2;

74 CCGAGAGTGCATCTCGGCAACCTGTGTGCTAGCCTAGTCTCTGGCTCCGACG 133


```
QY 923 GCAACTGAGGCTATTATTTATCTTTATCTGCTGAGGCTACGAAATGACAGAGCTCA 982
Db 661 GCAACTGAGGCTATTATTTATCTTTATCTGCTGAGGCTACGAAATGACAGAGCTCA 720
QY 983 AGCAGAGAGCAGAGGCAATTAAGGCAATGCTAAAAAGCTATTGAGAGCTCACTA 1042
Db 721 AGCAGAGAGCAGAGGCAATTAAGGCAATGCTAAAAAGCTATTGAGAGCTCACTA 780
QY 1043 CGCAGCTGAAACAGAGAGCAGAGAAATGGCCCTGATGAGATAGTTGCTGTGTC 1102
Db 781 CGCAGCTGAAACAGAGAGCAGAGAAATGGCCCTGATGAGATAGTTGCTGTGTC 840
QY 1103 ATTGAATTGTATAACCAATATTTGGTACGATCTTTAGCTGACATATTTTCCAT 1162
Db 841 ATTGAATTGTATAACCAATATTTGGTACGATCTTTAGCTGACATATTTTCCAT 900
QY 1163 AAGCAGATGTTGACCCATGCTGTTAAAAACAGACATTTGCCCATGTCGAATGTGAC 1222
Db 901 AAGCAGATGTTGACCCATGCTGTTAAAAACAGACATTTGCCCATGTCGAATGTGAC 960
QY 1223 ATCTCMAAGCTTTGGCAATGAGGTGATGTTGAAAGATGATCAAGTCTTTACAGTC 1282
Db 961 ATCTCMAAGCTTTGGCAATGAGGTGATGTTGAAAGATGATCAAGTCTTTACAGTC 1020
QY 1283 CCTGATCCATGAAATATCTATATAGTCCTCTCCCATGAGAGAAATTCGACGAG 1342
Db 1021 CCTGATCCATGAAATATCTATATAGTCCTCTCCCATGAGAGAAATTCGACGAG 1080
QY 1343 ACCGATATCTTGGAATATGCTTCACTACAGGAAACGATGAACGGCTCTGAGAGAAC 1402
Db 1081 ACCGATATCTTGGAATATGCTTCACTACAGGAAACGATGAACGGCTCTGAGAGAAC 1140
QY 1403 GTGAGTCACAAATGAATGAAGTCTACAGCTGTTAAACCATGACAAATTTCTGACAG 1462
Db 1141 GTGAGTCACAAATGAATGAAGTCTACAGCTGTTAAACCATGACAAATTTCTGACAG 1200
QY 1463 GATGTATTTCTCTCATGAGACCAACCTTGAAGAGAGAAATCTCTATCAAGAG 1522
Db 1201 GATGTATTTCTCTCATGAGACCAACCTTGAAGAGAGAAATCTCTATCAAGAG 1260
QY 1523 ACTGCTGTTGAGAAATTAATCT 1546
Db 1261 ACTGCTGTTGAGAAATTAATCT 1284

RESULT 6
AAS25884
ID AAS25884 standard; cDNA; 2306 BP.
XX
AC AAS25884;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human cDNA encoding a novel secreted protein, Seq ID 63.
XX
KW Human; immunosuppressive; antiarthritic; ss; antineoplastic;
KW cytosolic; cardiant; vasotropic; cerebroprotective; nootropic;
KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;
KW vulnary; secreted protein; rheumatoid arthritis;
KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;
KW cerebrovascular disorder; cerebral ischemia; angiogenesis;
KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;
KW corneal infection; wound healing; epithelial cell proliferation;
KW skin aging; food additive; preservative; antiproliferative.
XX
OS Homo sapiens.
XX
PN W0000155322-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01341.
XX
```

```
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-MAR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 05-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 06-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
```


PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 01-NOV-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-024617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 05-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-025678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 X1 Rosen CA, Barash SC, Ruben SM;
 XX
 DR P-PSDB; AAU15897.
 XX
 PT New nucleic acid molecules encoding 461 human secreted proteins for
 diagnosing, preventing, treating or ameliorating medical conditions and

PT used as food additives or preservatives -
 XX
 XX Claim 1; SEQ ID No 63; 980bp; English.
 XX
 CC The invention relates to isolated nucleic acid molecules and proteins are used to
 CC encoded secreted proteins. The nucleic acids and proteins are used to
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They
 CC are also used in diagnosing a pathological condition or susceptibility
 CC to a pathological condition. Antibodies to the proteins can also
 CC be used in alleviating symptoms associated with the disorders and in
 CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
 CC immunosorbent assays (ELISA). Disorders which are diagnosed or treated
 CC include autoimmune diseases e.g. rheumatoid arthritis,
 CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
 CC e.g. cerebral ischaemia, angioneitis, nervous system disorders e.g.
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi
 CC and ocular disorders e.g. corneal infection, and many other
 CC disorders listed in the specification. The polypeptides can also
 CC be used to aid wound healing and epithelial cell proliferation, to
 CC prevent skin aging due to sunburn, to maintain organs before
 CC transplantation, for supporting cell culture of primary tissues, to
 CC regenerate tissues and in chemotaxis. The polypeptides can also be used
 CC as a food additive or preservative to increase or decrease storage
 CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
 CC minerals, cofactors and other nutritional components. The present
 CC sequence encodes a novel secreted protein of the invention.
 Query Match 66.4%; Score 1178.8; DB 22; Length 2306;
 Best Local Similarity 99.0%; Pred. No. 0;
 Matches 1217; Conservative 0; Mismatches 8; Indels 4; Gaps 3;
 QY 547 CTGTAACCCGACACAGATTTCACGGTGGCCAGGTTGGGAGACCCGTGCAAGTCTC 606
 Db CTGTAACCCGACACAGATTTCACGGTGGCCAGGTTGGGAGACCCGTGCAAGTCTC 71
 QY 607 TTGGTTGGCCCTTCATCCACGCGCGGGCGGTGACCTTGGCGACAAATCCATCGGC 666
 Db TTGGTTGGCCCTTCATCCACGCGCGGGCGGTGACCTTGGCGACAAATCCATCGGC 131
 QY 72 TTGGTTGGCCCTTCATCCACGCGCGGGCGGTGACCTTGGCGACAAATCCATCGGC 131
 Db TTGGTTGGCCCTTCATCCACGCGCGGGCGGTGACCTTGGCGACAAATCCATCGGC 191
 QY 667 TTATGAGAGATGGGGGTCTGGAACCCGTCACTTTAACTCCCGGGACCCGGAATGAGT 726
 Db TTATGAGAGATGGGGGTCTGGAACCCGTCACTTTAACTCCCGGGACCCGGAATGAGT 191
 QY 727 CATGCCAGTGTCTACCCGGGTGAGTAGAATTGGCATCATGATGGCAATCGAA 786
 Db CATGCCAGTGTCTACCCGGGTGAGTAGAATTGGCATCATGATGGCAATCGAA 251
 QY 132 CATGCCAGTGTCTACCCGGGTGAGTAGAATTGGCATCATGATGGCAATCGAA 251
 Db CATGCCAGTGTCTACCCGGGTGAGTAGAATTGGCATCATGATGGCAATCGAA 846
 QY 787 AGGCGAATAATTCTGCATCTATTTCAGAGGCAATTCAGATGCTATGAAAT 846
 Db AGGCGAATAATTCTGCATCTATTTCAGAGGCAATTCAGATGCTATGAAAT 311
 QY 252 AGGCGAATAATTCTGCATCTATTTCAGAGGCAATTCAGATGCTATGAAAT 311
 Db AGGCGAATAATTCTGCATCTATTTCAGAGGCAATTCAGATGCTATGAAAT 905
 QY 847 AGGCGAATAATTCTGCATCTATTTCAGAGGCAATTCAGATGCTATGAAAT 905
 Db AGGCGAATAATTCTGCATCTATTTCAGAGGCAATTCAGATGCTATGAAAT 371
 QY 312 AGGCGAATAATTCTGCATCTATTTCAGAGGCAATTCAGATGCTATGAAAT 371
 Db AGGCGAATAATTCTGCATCTATTTCAGAGGCAATTCAGATGCTATGAAAT 965
 QY 906 TTTTATATTTAGCGCGGCACTGTGGGCTATTTTATCTTTTATCTGCGAAGCTAC 965
 Db TTTTATATTTAGCGCGGCACTGTGGGCTATTTTATCTTTTATCTGCGAAGCTAC 431
 QY 372 TTTTATATTTAGCGCGGCACTGTGGGCTATTTTATCTTTTATCTGCGAAGCTAC 431
 Db TTTTATATTTAGCGCGGCACTGTGGGCTATTTTATCTTTTATCTGCGAAGCTAC 1025
 QY 966 GGAATCGAAGAGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1025
 Db GGAATCGAAGAGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 491
 QY 432 GGAATCGAAGAGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 491
 Db GGAATCGAAGAGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1085
 QY 1026 TTGGAAGGCTTCAACTAGAGCACTGTAACAGAGAGAGAGAGAGAGAGAGAGAG 1085
 Db TTGGAAGGCTTCAACTAGAGCACTGTAACAGAGAGAGAGAGAGAGAGAGAGAG 551
 QY 492 TTGGAAGGCTTCAACTAGAGCACTGTAACAGAGAGAGAGAGAGAGAGAGAGAG 551
 Db TTGGAAGGCTTCAACTAGAGCACTGTAACAGAGAGAGAGAGAGAGAGAGAGAG 1145
 QY 1086 ATAGTTGCTGTGTGATGATTAATTTATATACCAATGATTTGGACGATCTTAAGT 1145
 Db ATAGTTGCTGTGTGATGATTAATTTATATACCAATGATTTGGACGATCTTAAGT 611

QY 1146 GCACCATATTTTCATTAAGACATGTGTGACCCATGGCTGTTAAACACAGACCTGCC 1205
 DB 612 GCACCATATTTTCATTAAGACATGTGTGACCCATGGCTGTTAAACACAGACCTGCC 671
 QY 1206 CCATGTGCAATGTGCATCTCAACAGCTTTGGGAATGAGGTGATGAAAGATGAT 1265
 DB 672 CCATGTGCAATGTGCATCTCAACAGCTTTGGGAATGAGGTGATGAAAGATGAT 731
 QY 1266 CAGGTCTTTACAGTCCCTGATCCATGAATATATATAGTGTCTCCCTCCCATGAG 1325
 DB 732 CAGGTCTTTACAGTCCCTGATCCATGAATATATATAGTGTCTCCCTCCCATGAG 791
 QY 1326 AGGATATCCAGCCGAGACCCGATCTCTGGAATGCTTCAATCAGGGAACGATGAC 1385
 DB 792 AGGATATCCAGCCGAGACCCGATCTCTGGAATGCTTCAATCAGGGAACGATGAC 851
 QY 1386 CGCCTCTGAGAGAACAGTGCAGTCAACAAATGAAGTCAAGCTGTGTAACCAATGAG 1445
 DB 852 CGCCTCTGAGAGAACAGTGCAGTCAACAAATGAAGTCAAGCTGTGTAACCAATGAG 911
 QY 1446 CAATCTCTGAGAGAGAGATCTATCTCTGATGACCAACCACTTTGAAAGAGACG 1505
 DB 912 CAATCTCTGAGAGAGAGATCTATCTCTGATGACCAACCACTTTGAAAGAGACG 971
 QY 1506 AAACCTCAATCAAGAGACTGCTGTGAGAAATTAATCTTAATCTGTGAATGAG 1565
 DB 972 AAACCTCAATCAAGAGACTGCTGTGAGAAATTAATCTTAATCTGTGTGAATGAG 1031
 QY 1566 AAACCTCAATCAATTAATTAACAACTCCGATCAGGAGCTTCTTCAATTAAT 1625
 DB 1032 AAACCTCAATCAATTAATTAACAACTCCGATCAGGAGCTTCTTCAATTAAT 1091
 QY 1626 TCGAATAATTTAAATTAAGATGATCTGAAATGCTGATGATTAATTAATGCT 1685
 DB 1092 TCGAATAATTTAAATTAAGATGATCTGAAATGCTGATGATTAATTAATGCT 1151
 QY 1686 ATAGTAAATTAATGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1745
 DB 1152 ATAGTAAATTAATGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1208
 QY 1746 ATATTTTCATTAAGAGAGTTTCTCTGAG 1774
 DB 1209 ATATTTTCATTAAGAGAGTTTCTCTGAG 1237

RESULT 7
 ABX73225
 ID ABX73225 standard; DNA; 2306 BP.
 XX
 AC ABX73225;
 XX
 DT 18-MAR-2003 (first entry)
 XX
 DE Human novel polynucleotide #53.
 XX
 KW Human; gene; ds; neural disorder; immune system disorder; renal disorder;
 KW muscular disorder; respiratory disease; reproductive disorder;
 KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
 KW hyperproliferative disorder; inflammatory disease; allergic reaction;
 KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
 KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
 KW haemostatic; antiarteriosclerotic.
 XX
 OS Homo sapiens.
 XX
 PN US200212753-A1.
 XX
 PD 19-SEP-2002.
 XX
 PR 17-JAN-2001; 2001US-0764864.
 XX
 PR 31-JAN-2000; 2000US-179065P.
 PR 04-FEB-2000; 2000US-180628P.

PR 28-JUN-2000; 2000US-214886P.
 PR 07-JUL-2000; 2000US-216647P.
 PR 07-JUL-2000; 2000US-216880P.
 PR 11-JUL-2000; 2000US-217487P.
 PR 11-JUL-2000; 2000US-217496P.
 PR 14-JUL-2000; 2000US-218290P.
 PR 26-JUL-2000; 2000US-220363P.
 PR 26-JUL-2000; 2000US-220964P.
 PR 14-AUG-2000; 2000US-224518P.
 PR 14-AUG-2000; 2000US-224519P.
 PR 14-AUG-2000; 2000US-22567P.
 PR 14-AUG-2000; 2000US-22568P.
 PR 14-AUG-2000; 2000US-22570P.
 PR 14-AUG-2000; 2000US-225447P.
 PR 14-AUG-2000; 2000US-225757P.
 PR 14-AUG-2000; 2000US-22578P.
 PR 22-AUG-2000; 2000US-226868P.
 PR 30-AUG-2000; 2000US-228924P.
 PR 01-SEP-2000; 2000US-229287P.
 PR 01-SEP-2000; 2000US-229343P.
 PR 01-SEP-2000; 2000US-229344P.
 PR 01-SEP-2000; 2000US-229345P.
 PR 05-SEP-2000; 2000US-229509P.
 PR 05-SEP-2000; 2000US-229513P.
 PR 08-SEP-2000; 2000US-231413P.
 PR 21-SEP-2000; 2000US-234223P.
 PR 21-SEP-2000; 2000US-234274P.
 PR 25-SEP-2000; 2000US-234997P.
 PR 27-SEP-2000; 2000US-235834P.
 PR 29-SEP-2000; 2000US-236327P.
 PR 29-SEP-2000; 2000US-236367P.
 PR 29-SEP-2000; 2000US-236368P.
 PR 29-SEP-2000; 2000US-236369P.
 PR 29-SEP-2000; 2000US-236370P.
 PR 02-OCT-2000; 2000US-236802P.
 PR 02-OCT-2000; 2000US-237037P.
 PR 02-OCT-2000; 2000US-237038P.
 PR 02-OCT-2000; 2000US-237039P.
 PR 02-OCT-2000; 2000US-237040P.
 PR 13-OCT-2000; 2000US-239935P.
 PR 20-OCT-2000; 2000US-240960P.
 PR 20-OCT-2000; 2000US-241785P.
 PR 20-OCT-2000; 2000US-241809P.
 PR 01-NOV-2000; 2000US-244617P.
 PR 17-NOV-2000; 2000US-249299P.
 PR 08-DEC-2000; 2000US-251856P.
 PR 08-DEC-2000; 2000US-251868P.
 PR 08-DEC-2000; 2000US-251869P.
 XX
 PA (ROSE/) ROSEN C A.
 PA (RUBS/) RUBEN S M.
 PA (BARA/) BARASH S C.
 XX
 PI Rosen CA, Ruben SM, Barash SC;
 XX
 DR WPI; 2003-14744/14.
 DR P-PSDB; ABUS4965.
 XX
 PT New polypeptides and nucleic acids, useful in gene therapy for
 PT treating, inhibiting or preventing e.g. neural, immune system,
 PT muscular, respiratory, reproductive, gastrointestinal, pulmonary,
 PT cardiovascular or renal disorders -
 XX
 PS Claim 1; SEQ ID NO 63; 402bp; English.
 XX
 CC The invention relates to human novel polypeptides and their associated
 CC polynucleotides. The polypeptides and polynucleotides are useful in gene
 CC therapy for treating, inhibiting or preventing neural disorders, immune
 CC system disorders (e.g. systemic lupus erythematosus, rheumatoid
 CC arthritis and multiple sclerosis), muscular disorders, respiratory
 CC diseases (e.g. nasal vestibulitis, nasal polyps and sinusitis),
 CC reproductive disorders, gastrointestinal disorders, pulmonary disorders,
 CC cardiovascular disorders (e.g. congenital heart defects, Ebstein's

anomaly and hypoplastic left heart syndrome), renal disorders (e.g. acute kidney failure and end-stage renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and leukemia), inflammatory diseases (e.g. septic shock, bursts and appendicitis), allergic reactions and conditions (e.g. asthma), blood related disorders (e.g. thrombosis, atherosclerosis and myocardial infarction) and cancerous diseases. Sequences ABX3173-ABX74167 represent human novel polynucleotides of the invention.

Sequence 2306 BP; 753 A; 391 C; 438 G; 724 T; 0 other;

Query Match 66.4%; Score 1178.8; DB 25; Length 2306;

Best Local Similarity 99.0%; Pred. No. 0;

Matches 1217; Conservative 0; Mismatches 8; Indels 4; Gaps 3;

547 CTGTACCCGACACGATATTCACGGTCCACGGTTGGGAGACCGCTGACCTC 606
 12 CTGTACCCGACACGATATTCACGGTCCACGGTTGGGAGACCGCTGACCTC 71
 607 TTGGTGGCCCTCATCAACGGGCGGGGCTCAGCTCGACACAGATCCATCTGC 666
 72 TTGGTGGCCCTCATCAACGGGCGGGGCTCAGCTCGACACAGATCCATCTGC 131
 667 TTATAGAGATGGGCTCTGAGCCCTCATCTTAACTTCCCGGACCCGCAATGAGCT 726
 132 TTATAGAGAGGGGCTCTGAGCCCTCATCTTAACTTCCCGGACCCGCAATGAGCT 191
 727 CATCCCATCTCAACCCGGGTGACATGATGATGATGATGATGATGATGATGAT 786
 192 CATCCCATCTCAACCCGGGTGACATGATGATGATGATGATGATGATGATGAT 251
 787 AGGCACAAAATTCGATATCTTCAAAAGGACATCAAGTCAATGATGATGAT 846
 252 AGGCACAAAATTCGATATCTTCAAAAGGACATCAAGTCAATGATGATGAT 311
 847 AGGCAAAATTCGATATCTTCAAAAGGACATCAAGTCAATGATGATGATGAT 905
 312 AGGCAAAATTCGATATCTTCAAAAGGACATCAAGTCAATGATGATGATGAT 371
 906 TTTTATATTAAGGGGACATCTGAGGCTATTTTCTTTATTCGCTCAAGGCTAC 965
 372 TTTTATATTAAGGGGACATCTGAGGCTATTTTCTTTATTCGCTCAAGGCTAC 431
 966 GGAATGCAAGGCTCAACGAGACGAGGACATTTAAGGACATGCTAAAGCTA 1025
 433 GGAATGCAAGGCTCAACGAGACGAGGACATTTAAGGACATGCTAAAGCTA 491
 1026 TTGGAAGGCTTCAACGAGACGAGGACATTTAAGGACATGCTAAAGCTA 1085
 492 TTGGAAGGCTTCAACGAGACGAGGACATTTAAGGACATGCTAAAGCTA 551
 1086 ATAGTGTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1145
 552 ATAGTGTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 611
 1146 GGAACCATATTTCCATAGACATGATGATGATGATGATGATGATGATGATGAT 1205
 612 GGAACCATATTTCCATAGACATGATGATGATGATGATGATGATGATGATGAT 671
 1206 CCATGTGCAATGTGACATATCAAAAGCTTTGGAAATGAGGTGATGATGAT 1265
 672 CCATGTGCAATGTGACATATCAAAAGCTTTGGAAATGAGGTGATGATGAT 731
 1266 CAGTGTCTTTACAAAGCCCTGTATGATGATGATGATGATGATGATGATGAT 1325
 732 CAGTGTCTTTACAAAGCCCTGTATGATGATGATGATGATGATGATGATGAT 791
 1326 AGGATATTCGAGGACCGGATATCTGATGATGATGATGATGATGATGATGAT 1385
 792 AGGATATTCGAGGACCGGATATCTGATGATGATGATGATGATGATGATGAT 851
 1386 CGCCTCGAGGAAACGTCAGTCAAAATGAATGAATGATGATGATGATGATGAT 1445

852 CGCCTCGAGGAAACGTCAGTCAAAATGAATGATGATGATGATGATGATGAT 911
 1446 CAATTCGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1505
 912 CAATTCGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 971
 1506 AAATCTCTAATCAAGACGCTGCTGAGGAAATTAATTTAAATCTGCTGATTA 1565
 972 AAATCTCTAATCAAGACGCTGCTGAGGAAATTTAAATCTGCTGATTAATGA 1031
 1566 AAATCTGAAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1625
 1032 AAATCTGAAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1091
 1626 TGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1685
 1092 TGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1151
 1686 ATAGTTAAATGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1745
 1152 ATAGTTAAATGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1208
 1746 ATATTTTCAATGAGGACATGATGATGATGATGATGATGATGATGATGAT 1774
 1209 ATATTTTCAATGAGGACATGATGATGATGATGATGATGATGATGATGAT 1237

RESULT 8
 ABL90796
 ID ABL90796 standard; cDNA; 1249 BP.
 XX
 AC ABL90796;
 XX
 DT 24-MAY-2002 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 1358.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antifungal;
 KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PD 29-NOV-2001.
 XX
 PF 18-MAY-2001; 2001WO-US16450.
 XX
 PR 19-MAY-2000; 2000US-205515P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Birse CE, Rosen CA;
 XX
 DR MPI; 2002-122018/16.
 DR P-FSDB; ABB90387.
 XX
 PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
 PT prevention of neural, immune system, muscular, reproductive,
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
 PT disorders -
 XX
 PS Claim 4; SEQ ID NO 1358; 2081bp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (ABL9443-ABL90853) and proteins
 CC (ABL89040-ABL90444) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast

CC and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the CC printed specification, but was obtained in electronic format directly CC from WIPO at ftp.wipo.int/pub/published_pot_sequences.
 CC
 XX

Sequence 1249 BP; 379 A; 258 C; 273 G; 335 T; 4 others;

Query Match 65.8%; Score 1167.2; DB 24; Length 1249;
 Best Local Similarity 98.7%; Pred. No. 1.3e-314;
 Matches 1206; Conservative 1; Mismatches 11; Indels 4; Gaps 3;

QY 547 CTGTACCCCGACACGAATTCAGCGTCCGACGCTTGGGAGACCGCTGCAATCTC 606
 DB 18 CTGTACCCCGACACGAATTCAGCGTCCGACGCTTGGGAGACCGCTGCAATCTC 77
 QY 607 TTGTTGGCCCTCATCCACGCGCGGGGCTGACCTTCGACAGAAATCCATCTGAC 666
 DB 78 TTGTTGGCCCTCATCCACGCGCGGGGCTGACCTTCGACAGAAATCCATCTGAC 137
 QY 667 TTATGAGAGAGAGGCGCTGAGAGCGGTCACTTTAACTCCCGGAGACCGCAATGAGT 726
 DB 138 TTATGAGAGAGAGGCGCTGAGAGCGGTCACTTTAACTCCCGGAGACCGCAATGAGT 197
 QY 727 CATCCCATGCTCACCCCGGCTGAGTGAATGTTGCAATGATCGGCAATCTGAA 786
 DB 198 CATCCCATGCTCACCCCGGCTGAGTGAATGTTGCAATGATCGGCAATCTGAA 257
 QY 787 AGGCACAAAATTCGATCATCTATCAAGGCAATGAGTGAATGCTAGAGT 846
 DB 258 AGGCACAAAATTCGATCATCTATCAAGGCAATGAGTGAATGCTAGAGT 317
 QY 847 AGGCAAAAAATGCGCCCTGGGTGATCACTATTCATTTTTCTTCTGCTCTT 906
 DB 318 AGGCAAAAAATGCGCCCTGGGTGATCACTATTCATTTTTCTTCTGCTCTT 377
 QY 907 TTTATTTAATGAGGCGGCACTGCGGCTATTTATCTTTATCTGCTGAGAGCTG 966
 DB 378 TTTATTTAATGAGGCGGCACTGCGGCTATTTATCTTTATCTGCTGAGAGCTG 437
 QY 967 GAATGCAAGACTCTCAAGCAGAGCAGAGCAATTTAAAGCAATGCTAAATAGCTAT 1026
 DB 438 GAATGCAAGACTCTCAAGCAGAGCAGAGCAATTTAAAGCAATGCTAAATAGCTAT 497
 QY 1027 TGGAGGCTTCACTGACGACACTGAACAGAGCAAGG-AAATGGGCGCTGATGAG 1085
 DB 498 TGGAGGCTTCACTGACGACACTGAACAGAGCAAGG-AAATGGGCGCTGATGAG 557
 QY 1086 ATAGTTGCTGTGTGATGATTAATTAACCAATGATTTGGTACGATCTTAACGT 1145
 DB 558 ATAGTTGCTGTGTGATGATTAATTAACCAATGATTTGGTACGATCTTAACGT 617
 QY 1146 GGAACCATATTTTCCATAGACAGATGTTGACCGAGGCTTAAACAGCAACTTGC 1205
 DB 618 GGAACCATATTTTCCATAGACAGATGTTGACCGAGGCTTAAACAGCAACTTGC 677
 QY 1206 CCATGTGCAAAATGTGACATACCTCAAGCTTTGGAAATGAGTGTGATGATGAT 1265
 DB 678 CCATGTGCAAAATGTGACATACCTCAAGCTTTGGAAATGAGTGTGATGATGAT 737
 QY 1266 CAGTGTCTTTACAGTCCCTGATCCATGAAATATCTATGCTCCCTCCCATGAG 1325
 DB 738 CAGTGTCTTTACAGTCCCTGATCCATGAAATATCTATGCTCCCTCCCATGAG 797
 QY 1326 AGGATTAATCGAGGAGACCGCATGATCTGATTTGCTTCACTACAGGAAACGATGAC 1385

DB 798 AGGATTAATCGAGGAGACCGCATGATCTGATTAATGCTTCACTACAGGAAACGATGAC 857
 QY 1386 CGCTCTGAGGAAACAGCTGACATCAAAATGAAAGTCAAGCTGTAAACCTGAG 1445
 DB 888 CGCTCTGAGGAAACAGCTGACATCAAAATGAAAGTCAAGCTGTAAACCTGAG 917
 QY 1446 CAAATTCGTGCGAGTGAATGTTATTCCTGATGTAACCAACCACTTTGAAAGAGCG 1505
 DB 918 CAAATTCGTGCGAGTGAATGTTATTCCTGATGTAACCAACCACTTTGAAAGAGCG 977
 QY 1506 AAATCTCTATCAAGACCTGCTGTCGAGAAATTAATCTTAAATCTGTAAATGAA 1565
 DB 978 AAATCTCTATCAAGACCTGCTGTCGAGAAATTAATCTTAAATCTGTAAATGAA 1637
 QY 1566 AAATCTCTATCAAGACCTGCTGTCGAGAAATTAATCTTAAATCTGTAAATGAA 1625
 DB 1038 AAATCTCTATCAAGACCTGCTGTCGAGAAATTAATCTTAAATCTGTAAATGAA 1697
 QY 1626 TGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1685
 DB 1098 TGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1157
 QY 1686 ATAGTTAAATGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1745
 DB 1158 ATAGTTAAATGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1214
 QY 1746 ATATTTTCAATAGGCAATTTTC 1767
 DB 1215 ATATTTTCAATAGGCAATTTTC 1236

RESULT 9

AA172094
 ID AA172094 standard; cDNA, 2145 BP.

AC AA172094;
 DT 25-MAR-2002 (first entry)

XX Mouse GRAIL cDNA.

XX Murine; human; GRAIL; anergy; attenuation; tyrosine; phosphorylation;
 XX antigenic stimulation; interleukin-2; gene therapy; polymorphism; IL-2;
 XX autoimmune disease; tumour cell; cancer; transplant rejection; ss.

OS Mus musculus.

XX Key Location/Qualifiers

XX CDS 358..1641

XX FT /*tag= a

XX FT /product= "GRAIL"

XX W0200165943-A1.

XX PD 15-NOV-2001.

XX PE 11-MAY-2001; 2001MO-US15385.

XX PR 11-MAY-2000; 2000US-203513P.

XX (STRD) UNIV LELAND STANFORD JUNIOR.

XX FI Ford GS, Bloom D, Fatman CG;

XX DR WPI; 2002-055597/07.

XX DR P-FSDB; AAB47799.

XX Novel nucleic acid sequences of anergy associated genes, including
 XX GRAIL gene useful in the evaluation of pathophysiology or immunotherapy
 XX of cancer, autoimmune disease and transplant rejection -

XX Claim 2; Page 42-44; 50pp; English.

XX
AC AAs62340;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human cDNA encoding a novel secreted protein, Seq ID 519.
XX
KW Human; immunosuppressive; antiarthritic; ss; antirheumatic;
KW cytosolic; cardiant; vasotropic; cerebroprotective; nootropic;
KW neuroprotective; antibacterial; virucide; fungicide; ophthalmologic;
KW unitary; secreted protein; rheumatoid arthritis;
KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;
KW cerebrovascular disorder; cerebral ischemia; angiogenesis;
KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;
KW corneal infection; wound healing; epithelial cell proliferation;
KW skin aging; food additive; preservative; antiproliferative.
XX
OS Homo sapiens.
XX
PN W0200155322-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01341.
XX
PE 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 07-JUL-2000; 2000US-0216547.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0228287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.

PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231966.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234222.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235634.
PR 27-SEP-2000; 2000US-0235636.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 13-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.

PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-488783/53.
 DR P-PSDB; AAV16353.
 XX
 PT New nucleic acid molecules encoding 461 human secreted proteins for
 PT diagnosing, preventing, treating or ameliorating medical conditions and
 PT used as food additives or preservatives -
 XX
 PS Claim 1; SEQ ID No 519; 980bp; English.
 XX
 CC The invention relates to isolated nucleic acid molecules and their
 CC encoded secreted proteins. The nucleic acids and proteins are used to
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They
 CC are also used in diagnosing a pathological condition or susceptibility
 CC to a pathological condition. Antibodies to the proteins can also
 CC be used in alleviating symptoms associated with the disorders and in
 CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
 CC immunoassay assays (ELISA). Disorders which are diagnosed or treated
 CC include autoimmune diseases e.g. rheumatoid arthritis,
 CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
 CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi
 CC and ocular disorders e.g. corneal infection, and many other
 CC disorders listed in the specification. The polypeptides can also
 CC be used to aid wound healing and epithelial cell proliferation, to
 CC prevent skin aging due to sunburn, to maintain organs before
 CC transplantation, for supporting cell culture of primary tissues, to
 CC regenerate tissues and in chemotaxis. The polypeptides can also be used
 CC as a food additive or preservative to increase or decrease storage
 CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
 CC minerals, cofactors and other nutritional components. The present
 CC sequence encodes a novel secreted protein of the invention.
 Query Match 65.3%; Score 1157.6; DB 22; Length 1250;
 Best Local Similarity 98.8%; Pred. No. 5e-312;
 Matches 1208; Conservative 0; Mismatches 10; Indels 5; Gaps 4;
 QY 547 CTGTAAACCCGACACGAATTTCAACGCTCCACAGGTTTGGGAGACACCGTGAAGTCTC 606
 DB 18 CTGTAAACCCGACACGAATTTCAACGCTCCACAGGTTTGGGAGACACCGTGAAGTCTC 77
 QY 607 TTGGTGGCCCTCATTCACACGCGGCGGCGGCTGACCTTTCGAGCAAGATCATTTGGC 666
 DB 78 TTGGTGGCCCTCATTCACACGCGGCGGCGGCTGACCTTTCGAGCAAGATCATTTGGC 137
 QY 667 TTATAGAGATGGGCGCTGAGCGGTATCTTAACTTCCCGGACCCGCAATGAGGT 726
 DB 138 TTATAGAGAGGCGCGCTGAGCGGTATCTTAACTTCCCGGACCCGCAATGAGGT 197
 QY 727 CATCCCATGCTTACCCGCGTGCATGATGATTTGCAATCAATGATGGCATCTGAA 786
 DB 198 CATCCCATGCTTACCCGCGTGCATGATGATTTGCAATCAATGATGGCATCTGAA 257

QY 787 AGGCACAAAATTCTGCATCTATTTCAAGAGGCACTACAGTGCATAGTGTATAGAGT 846
 DB 258 AGGCACAAAATTCTGCATCTATTTCAAGAGGCACTACAGTGCATAGTGTATAGAGT 317
 QY 847 AGGCAAAAACATGAGCCCTTGGGAGTACATCAATCAA-TTTTTCGTTCTGTGTCTCT 905
 DB 318 AGGCAAAAACATGAGCCCTTGGGAGTACATCAATCAA-TTTTTCGTTCTGTGTCTCT 377
 QY 906 TTTTATATATACGGCGGCAATCTGTGGCTATTTTATCTTTATCTGCTCGAAGGCTAC 965
 DB 378 TTTTATATATACGGCGGCAATCTGTGGCTATTTTATCTTTATCTGCTCGAAGGCTAC 437
 QY 966 GGAATGCAAGACTCAAGCAGAGCAAGGCAATTTAAAGCAGATGCTAAAAAGCTA 1025
 DB 438 GGAATGCAAGACTCAAGCAGAGCAAGGCAATTTAAAGCAGATGCTAAAAAGCTA 497
 QY 1026 TTGAAAGCTTCACTACGCACTGTAACAAGAGACAAAG-AAATTTGCTCTGATGCA 1084
 DB 498 TTGAAAGCTTCACTACGCACTGTAACAAGAGACAAAGAAATTTGCTCTGATGCA 557
 QY 1085 GATAGTTGCTGTGTCATGATGATTTGTAACCAATGATTTGGTACGATCTTAAAG 1144
 DB 558 GATAGTTGCTGTGTCATGATGATTTGTAACCAATGATTTGGTACGATCTTAAAG 617
 QY 1145 TGCACCAATTTTCCATTAAGACATGTTGACCCATGCTGTTAAACCAAGACTTGC 1204
 DB 618 TGCACCAATTTTCCATTAAGACATGTTGACCCATGCTGTTAAACCAAGACTTGC 677
 QY 1205 CCGATGCGAAATATGACATATCTCAAGGCTTTGGAGATTTAGGTGTTGAAGTGA 1264
 DB 678 CCGATGCGAAATATGACATATCTCAAGGCTTTGGAGATTTAGGTGTTGAAGTGA 737
 QY 1265 TCACTGTCTTCAAGCTCCCTGTTATTCATGAATTAATTAATGCTCTCTCCATGA 1334
 DB 738 TCACTGTCTTCAAGCTCCCTGTTATTCATGAATTAATTAATGCTCTCTCCATGA 797
 QY 1335 GAGATATATGACAGAGACACCATATCATGATGATTTGCTCAATGACAGGAACAATGA 1384
 DB 798 GAGATATATGACAGAGACACCATATCATGATGATTTGCTCAATGACAGGAACAATGA 857
 QY 1385 CCGCTCTGAGAGAACAGTGCACATCAATGAAGTCTTACGCTGTTAAACCTGA 1444
 DB 858 CCGCTCTGAGAGAACAGTGCACATCAATGAAGTCTTACGCTGTTAAACCTGA 917
 QY 1445 GCAATTCGTGGCGATGATGTTATTCATGATGACCAACCACTTGAAGAGAC 1504
 DB 918 GCAATTCGTGGCGATGATGTTATTCATGATGACCAACCACTTGAAGAGAC 977
 QY 1505 GAAACTCTTATCAAGAGATGCTGTTCGAGAAATTAATCTTAAATCTGTGTAATAG 1564
 DB 978 GAAACTCTTATCAAGAGATGCTGTTCGAGAAATTAATCTTAAATCTGTGTAATAG 1037
 QY 1565 AAAACTGAAACATTAATTAACAGAACTGCACATCAAGGCGCTAGTTCTTAAATAA 1624
 DB 1038 AAAACTGAAACATTAATTAACAGAACTGCACATCAAGGCGCTAGTTCTTAAATAA 1097
 QY 1625 TTGATTAATTTAATTAATTAAGATGATCTGAAGTGTCTCAATGATATTAATAGC 1684
 DB 1099 TTGATTAATTTAATTAATTAAGATGATCTGAAGTGTCTCAATGATATTAATAGC 1157
 QY 1686 TATGTTTAAATGCTTAAATTAATTAATCTGTAACTTTTTCGACCAATCTCAATAT 1744
 DB 1158 TATGTTTAAAT-TGGCTTAAATTTTAACTGTATTAATCTTTTTCGACCAATCTCAATAT 1214
 QY 1745 AATAATTTTCAATAGGCAAGTTTC 1767
 DB 1215 AATAATTTTCAATAGGCAAGTTTC 1237

RESULT 11
 ABX73681
 ID ABX73681 standard; DNA; 1250 BP.


```

Db      498 TTGGAAGGCTTCAATTCAGCACTGAAACAGAGGACAAAGAAATTTGGCCCTGATGGA 557
QY      1085 GATAGTTGCTGCTGTGCTGATTAATGTATATAACCAATGATTTGTAGCATCTTAACG 1144
Db      558 GATAGTTGCTGCTGTGCTGATTAATGTATATAACCAATGATTTGTAGCATCTTAACG 617
QY      1145 TGCACCAATATTTTCCATTAAGACATGTGTGACCCATGGCTGTAAACAGAGCTTGC 1204
Db      618 TGCACCAATATTTTCCATTAAGACATGTGTGACCCATGGCTGTAAACAGAGCTTGC 677
QY      1205 CCATGTGCAAAATGTACATCTCAAGCTTTGGAAATGAGTGTGATGTGATGGA 1264
Db      678 CCATGTGCAAAATGTGACATCTCAAGCTTTGGAAATGAGTGTGATGTGATGGA 737
QY      1265 TCAGTCTTTCAAGTCCCTGTATCCATGAATTAATCTATAGTCCCTCCCATGAA 1324
Db      738 TCAGTCTTTCAAGTCCCTGTATCCATGAATTAATCTATAGTCCCTCCCATGAA 797
QY      1325 GAGGATATCGCAGCAGACCGCATCTGATATGCTTCAGTACAGGAAACAGATGA 1384
Db      798 GAGGATATCGCAGCAGACCGCATCTGATATGCTTCAGTACAGGAAACAGATGA 857
QY      1385 CCGCCTCTGAGAGAACAGTGCACATCAAAATGAAAGTCTGACGCTGTAAACCATGA 1444
Db      858 CCGCCTCTGAGAGAACAGTGCACATCAAAATGAAAGTCTGACGCTGTAAACCATGA 917
QY      1445 GCAATCTGTGCGAGTGTATATCTCTCAATGTCACCAACCTTTGAGAGAC 1504
Db      918 GCAATCTGTGCGAGTGTATATCTCTCAATGTCACCAACCTTTGAGAGAC 977
QY      1505 GAACTCTTATACAGACATCTCTTCAAGAAATTAATCTTAAATCTGTGTAATG 1564
Db      978 GAACTCTTATACAGACATCTCTTCAAGAAATTAATCTTAAATCTGTGTAATG 1037
QY      1565 AAACCTTGAACCATTAATATACAGACATCTGCAATCAAGGCTGATGCTTATATA 1624
Db      1038 AAACCTTGAACCATTAATATACAGACATCTGCAATCAAGGCTGATGCTTATATA 1097
QY      1625 TTGATTAATTAATTAATTAAGATGATCTGAAGTCTCAATGATTAATTAATG 1684
Db      1098 TTGATTAATTAATTAATTAAGATGATCTGAAGTCTCAATGATTAATTAATG 1157
QY      1685 TATAGTAAATGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1744
Db      1158 TATAGTAAATGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1214
QY      1745 AATATTTTCAAGCAGATTTC 1767
Db      1215 AATATTTTCAAGCAGATTTC 1237

RESULT 12
ABK12983 standard; cDNA, 1287 BP.
XX      ABK12983;
XX      AC
XX      ABK12983;
XX      23-APR-2002 (first entry)
XX      DE
XX      Mouse golliach cDNA sequence.
XX      Mouse; golliach protein; antiangiogenic; vasotropic; gene therapy;
XX      dosage form; angiogenesis; neurogenesis; tumour; vascularisation;
XX      cancer; ischaemia; neuroblastoma; neurodegenerative disease;
XX      unrecovered nerve trauma; gene; ss.
XX      OS
XX      Mus sp.
XX      FH
XX      Key
XX      CDS
XX      1..1287
XX      /product= "Mouse_golliach_protein"
XX      PT

```

```

XX      PN
XX      MO200193681-A1.
XX      PD
XX      13-DEC-2001.
XX      PF
XX      01-UN-2001; 2001WO-0818000.
XX      PR
XX      02-UN-2000; 2000US-0586398.
XX      PA
XX      (RBC ) UNITV CALIFORNIA.
XX      PI
XX      Harland R. Baker JC;
XX      WPI, 2002-147637/19.
XX      DR
XX      F-PSDB; AAD74918.
XX      PT
XX      New compositions comprising golliach proteins, useful for modulating
XX      angiogenesis or neurogenesis in mammals e.g. for preventing or
XX      treating undesirable vascularisation of a tumour, ischaemia or
XX      neurodegenerative disease
XX      Examples; Page 38; 45pp; English.
XX      CC
XX      The present invention relates to a new pharmaceutical composition that
XX      comprises a golliach polypeptide in dosage form. The golliach polypeptide
XX      has a sequence identity of at least 75% to the protein sequences
XX      (AAU74918-AAU74921) fully defined in the specification. The composition
XX      is useful for modulating angiogenesis or neurogenesis in mammals,
XX      particularly in humans or mice. Specifically, the composition is useful
XX      for the prophylactic and/or therapeutic treatment of excess angiogenesis
XX      e.g. undesirable vascularisation of a tumour or insufficient angiogenesis
XX      e.g. ischaemia, excess neurogenesis e.g. neuroblastoma or insufficient
XX      neurogenesis e.g. neurodegenerative disease or unrecovered nerve trauma.
XX      The present nucleic acid sequence encodes the mouse golliach protein
XX      that was used in the invention for modulating angiogenesis or
XX      neurogenesis.
XX      SO
XX      Sequence 1287 BP; 308 A; 320 C; 359 G; 300 T; 0 other;

Query Match      62.4%; Score 1106.8; DB 24; Length 1287;
Best Local Similarity 91.2%; Pred. No. 7.4e-298;
Matches 1174; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

QY      263 ATGGGAGCGCGCCGCGGAGCGGAGTCTCTGCGCGGCTGCGCTTTTCCAGATTG 322
Db      1 ATGGGAGCGCGCGCGGAGCGGAGTCTCTGCGCGGCTGCGCTTTTCCAGATTG 60
QY      323 CTGCAATGATGCTTCTCTGCTGCTGATCCGAGGACACCGGTTCCCGGAGGCTGAA 382
Db      61 CTGCAATGATGCTTCTCTGCTGCTGATCCGAGGACACCGGTTCCCGGAGGCTGAA 120
QY      383 GCAATGTGACCGCGTACCTCAAGTGTCTGAGCGGAGTTCGACACAGGAGTGAACGT 442
Db      121 GCGGTGTGACCGCGTACCTCAAGTGTCTGAGCGGAGTTCGACACAGGAGTGAACGT 180
QY      443 ACGGTGTGAGAGTGAAGAGAGAGGAGTGTGAGCGGAGTTCGACACAGGAGTGAACGT 502
Db      181 ACGGTGTGAGAGTGAAGAGAGAGGAGTGTGAGCGGAGTTCGACACAGGAGTGAACGT 240
QY      503 GCTGGAGTCTCTGATACCGCGCGGAGCGGAGTTCGATTAAGCTTGAACCGGACAG 562
Db      241 TCGGAGTCTCTGATACCGCGCGGAGCGGAGTTCGATTAAGCTTGAACCGGACAG 300
QY      563 AATTTCAGCGTGTCCAGCGGTTTGGGAGAGACCGGTGAAAGTCTTGTGTTGGCTTCATC 622
Db      301 AATTTCAGCGTGTCCAGCGGTTTGGGAGAGACCGGTGAAAGTCTTGTGTTGGCTTCATC 360
QY      623 CAAGCGGCGGAGGAGTGAACCTTGCAGACAAAGATCAATCTGCTTATGAAGATGAGCG 682
Db      361 CAAGCGGAGGAGTGAACCTTGCAGACAAAGATCAATCTGCTTATGAAGATGAGCG 420
QY      683 TCTGAGCGGTCAATCTTAACTTCCCGGAGACCCGATGAGGTCAATCCCATGTTCAC 742

```

Db	421	CTGGAGCGGCTAATCTTAACTTACCTCCGAGGACCCGGAATGAGGTCATCCCATCTGTCCAC	498
Qy	743	CCGGGTGCAGTACACATTGTGGCAATCAATATGGCAATCTGAAAGGCAACAAATTTCTG	802
Db	481	CCGGGTCTCGGGACATGTTGGCAATCATATTTGGCAATCTGAAAGGACAAAAATTTCTG	540
Qy	803	CAATCTATTCAAGAGGCAATCAAGTACCAATGGTCAATGAAATGGGAAAAAACAATGGC	862
Db	541	CAGTCTATTCAAGAGGCAATCCAGTCAATGTCTCTGAAAGTGGGAAAAAACAATGGC	600
Qy	863	CCTTGGGTGAATCACTATTCAATTTTTTGTGTGTGTGTCTGTCTTTTATATTAACGGCG	922
Db	601	CCTTGGGTGAATCATATTCAATTTTTTGTGTGTGTGTCTGTCTTTTATATTAACGGCG	660
Qy	923	GCAACTGAGGCTAATTTATCTTTTATCTGTCTGAAAGGCTACGAAATGGCAAGCTCA	982
Db	661	GCAACGTGGGCTAATTTATCTTTTATCTGTCTGAAAGTACGAAATGGCAAGCTCA	720
Qy	983	AGCAGAGACAGAGGCAATTAAGCGAGTCTAAAAAAGCTATTGAGAGCTTCACTA	1042
Db	721	AGCAGAGAGCAGAGGCAATTAAGCGAGTCTAAAAAAGCTATTGAGAGCTTCACTA	780
Qy	1043	CGCACACTGAAAACAAGAGACAGAGAAATTTGGCCCTGATGAGAGATGTTGTGTGTGC	1102
Db	781	CGCACCTTTGAAAACAAGAGACAGAGAAATTTGGCCCTGATGAGAGATGTTGTGTGTGC	840
Qy	1103	ATTGATTTGATTAACCAATTAATTTGGTACGATCTTAAGTGCACATATTTCAT	1162
Db	841	ATTGAGGCTCTAATGAGCAATTAATTTGGTACGATCTTAAGTGCACATATTTCAT	900
Qy	1163	AAGACATGTGTGACCATGAGCTGTAAAAACAAGACTTGCCCATGTGCATATGTGAC	1222
Db	901	AAGACATGTGTGACCATGAGCTGTAAAAACAAGACTTGCCCATGTGCATATGTGAC	960
Qy	1223	ATACTCAAACTTTTGGGAATTAAGTGTGATGTTGAATGATCAAGTCTTTAACAAGC	1282
Db	961	ATTCTCAAACTTTTGGGAATTAAGTGTGATGTTGAATGATCAAGTCTTTAACAAGT	1020
Qy	1283	CCTGATTCACATGAATATCTAATAGTCCCTCTCCCATGAAGAGATATCGACAGAG	1342
Db	1021	CCTGATTCATATGAACATCTAATATAGTCCCTCTCCCATGAAGAGATATCGACAGAG	1080
Qy	1343	ACCGCATATCTGATATAGCTTCACTAACAAGGAAACAGATGAACCGCTCGAGGAAACAC	1402
Db	1081	ACTGCATATCTGATATAGCTTCACTAACAAGGAAACAGATGAACCGCTCGAGGAAACAT	1142
Qy	1403	GTGCAGTCAACAAATGAAGTCAAGCTGTGAACCATGAAGCAATTTGTGGCAGTG	1462
Db	1141	GGCGAGTCAAGCAATGAATATCTAAGCTGTGAACCATGAAGCAATTTGTGGCAGTG	1200
Qy	1463	GATGTTATTCCTCATGATGACAAACCACTTTGAAAGAGAGAAATCCTAATCAAGAG	1522
Db	1201	GATGTTATTCCTCATGATGACAAACCACTTTGAAAGAGAGAAATCCTAATCAAGAG	1260
Qy	1523	ACTGCTGTTCGAGAAATTAATCTTAA	1549
Db	1261	GCAAGCTGTTGGGAGATTAATCTTAA	1287
RESULT 13			
ABK12992			
ID	ABK12992 standard; DNA, 1287 BP.		
XX	ABK12992;		
DT	23-APR-2002 (first entry)		
XX	Synthetic goliath DNA sequence #6.		
DE			
XX	Goliath protein; antiangiogenic; vasotropic; gene therapy;		
KM	dosage form; angiogenesis; neurogenesis; tumour; vasculatation;		
KM	cancer; ischaemia; neuroblastoma; neurodegenerative disease;		
KW	unrecovered nerve trauma; ds.		

XX	Synthetic.
XX	
XX	WO200193681-A1.
XX	
PX	13-DEC-2001.
PD	
PF	01-JUN-2001; 2001WC-US18000.
PR	02-JUN-2000; 2000US-0586398.
XA	(REGC) UNITV CALIFORNIA.
XA	
PA	Harland R, Baker JC;
PI	WPI; 2002-147637/19.
XX	
PT	New compositions comprising goliath proteins, useful for modulating
PT	angiogenesis or neurogenesis in mammals e.g. for preventing or
PT	treating undesirable vascularisation of a tumour, ischaemia or
PT	neurodegenerative disease -
PS	
XX	Disclosure; Page 42; 45pp; English.
XX	
CC	The present invention relates to a new pharmaceutical composition that
CC	comprises a goliath polypeptide in dosage form. The goliath polypeptide
CC	has a sequence identity of at least 75% to the protein sequences
CC	(AMU74918-AMU74921) fully defined in the specification. The composition
CC	is useful for modulating angiogenesis or neurogenesis in mammals,
CC	particularly in humans or mice. Specifically, the composition is useful
CC	for the prophylactic and/or therapeutic treatment of excess angiogenesis
CC	e.g. undesirable vascularisation of a tumour or insufficient angiogenesis
CC	e.g. ischaemia, excess neurogenesis e.g. neuroblastoma or insufficient
CC	neurogenesis e.g. neurodegenerative disease or unrecovers nerve trauma.
CC	The present nucleic acid sequence represents synthetic DNA sequence #6
CC	that is one of several artificial goliath nucleic acids (ABK12987-
CC	ABK12992) used in the invention for modulating angiogenesis or
CC	neurogenesis.
CC	
XX	
XX	Sequence 1287 BP; 312 A; 321 C; 357 G; 297 T; 0 other;
XX	
XX	Query Match 59.3%; Score 1052.4; DB 24; Length 1287;
XX	Best Local Similarity 88.6%; Pred.No.1.le-282;
XX	Matches 1140; Conservative 0; Mismatches 147; Indels 0; Gaps 0.
QY	263 ATGGGCGCCGCCCTCGGGCCGGGGGTCCCTGCAGCGGTGTGTGGTCTTCACAAATG 322
DB	1 ATGGGCGCGCGCCCGGATCGGGGTCACTGCAGCGGGGTGTGGAGCTCCCACTA 60
QY	323 CTGGCATGTGTCTTCTTGCTGTGCCCTGATGCCAGGCACCCTGTTCCCGGAGCTGA 382
DB	61 CTGGCTTGTGTCTTCTTCTTGCTGTGTGTGCCACGCGCCGGGTCCCGGAGCCGAA 120
QY	383 GAGATGTGACCGCGTAACCTGACGTGCTCGCGGGTTCGSCACAAGGAGTGAACGCT 442
DB	121 GCCGTGTAAACCGGTACTTCAACCTTCCGCGGGATTCCGACACCGAGTGAACCCG 180
QY	443 AGCGTGTGGAGTGAAGCAGGAGGAGCGCTGTAGCGCACAGACTGCGCTGTGAGCTGTG 502
DB	181 ACAGTGTGGAGTGAAGCAGGAGGAGCGCTGTAGCGCACAGCTCCGCTGTGAGCCGCTC 240
QY	503 GCTGGAGTTCGTGTATCCGCGCGAGCGGGCGCGGGGCGCTTAAAGCTGTAAACCGCAACG 562
DB	241 TCAGGAGTTCGTGTATCCGCGCGAGCGGGCGCGGGGCGCTTAAAGCTGTAAACCGCAAC 300
QY	563 AATTTCACGATGCCAACCGGTTTTGGGAGAGCACCGGTGCAAGTCTCTTGATGTGCCCCCTATC 622
DB	301 AATTTCACGATGCCAACCGGTTTTGGGAGAGCACCGGTGCAAGTCTCTTGATGTGCCCCCTATC 360
QY	623 CAACCGCGCGGGGGCGCATCTTGCACAAGAATCCAACTGACTTATGAGAGATGAGCG 682
DB	361 CATCGCGGTGAGGCTGTACCTTCGCGGACAAATCTCACTGTGCTTCAAGAGAGAGCT 420

QY	683	TCGAGAGCGCAACCTTAACTCCCGGGAACCGGCATGAGGTATCCCATCTCCAC	742
Db	421	TCGAGAGCGCACTTTTAACTCCCGGGAACCGGCATGAGGTATCCCATCTCCAC	480
QY	743	CCGGGTGCGATGACATTTGTGCATCATGATCCGCAATCGAATGGAAGCGACAAATTTCTG	802
Db	481	CCGGGTGCGAGGAGACATTTGTGCATCATGATGATGCGATCTGAAACGAAACAAATTTCTG	540
QY	803	CAATCATTTCAAAAGGCAATAGTGAAGATGATGAGTACGGGAAAAACATATGCG	862
Db	541	CAATCATTTCAAAAGGCAATAGTGAAGATGATGAGTACGGGAAAAACATATGCG	600
QY	923	GCAACGTGAGGCTATTTTATCTTTATCTTATCTGCTCGAAGGCTAAGAAATGGAAGAGCTCA	982
Db	661	GCAACGTGAGGCTATTTATCTTTATCTTATCTGCTCGAAGGCTAAGAAATGGAAGAGCTCA	720
QY	983	AGCAGGAAAGCAGAGGCAATTTAAAGCGATGCTAAAAAAAGCTATGAGAGGCTCACTA	1042
Db	721	AGTAGGAAAGCAGAGGCAATTTTAAAGCGATGCTAAAAAAAGCTATGAGAGGCTCAACTA	780
QY	1043	CGACACAGAAACAGAGGCAAGGAAATTTGGCCTGATGAGAGATGTGGCTGTGCTG	1102
Db	781	CGACCTTGAACATGAGGCAAGGAAATTTGGCCTGATGAGAGATGTGGCTGTGCTG	840
QY	1103	ATTGAATGTATAAACCAATGATTTGGTACGACATCTTAACGTGCAACCAATTTTCAT	1162
Db	841	ATTGACATATTTAAAGCCAAATATTTTGTGCGATCTTAACGTGCAACCAATTTTCAT	900
QY	1163	AATGACATGTGACCCATGAGCTGTAAACAAGACCTTGCCCATGCGCAATGTGAC	1222
Db	901	AATGACATGTGACCCGATGAGCTGTAAACAAGACCTTGCCCATGCGCAATGTGAC	960
QY	1223	ATACTAAAGCTTTGGGAAATGAGGTGATGTTGAAGATGATGATGATCTTTTCAAGTC	1282
Db	961	ATTCTAAAGCTCTGGGAAATGAGGTGATGTTGAAGATGATGATGATGATCTTTTCAAGTT	1020
QY	1283	CCGTATTCATATGAAATATATATATATGAGCTCCCTCCCATGAGAGATATATGAGCGAG	1342
Db	1021	CCGTATTCATATGAAATATATATATATGAGCTCCCTCCCATGAGAGATATATGAGCGAG	1080
QY	1343	ACCGATCATCTGGATATGCTTCAGTACAGGAGCAAGATGAACCGCTCTGAGAGAAC	1402
Db	1081	ACCGATCATCTGGATATGCTTCAGTACAGGAGCAAGGAGCAACCTCTGAGAGAACAT	1140
QY	1403	GTCGATTCACAAATGAAGTCTACAGCTGGTAAACATGAAGCAATTCGTGGACAGT	1462
Db	1141	GTCGATTCAGCAATGAAGTCTACAGCTGGTAAACATGAAGCAATTCGTGGACAGT	1200
QY	1463	GATGTATTCCTCATGATGACACCAACCTTTGAAGAGCAAACTCTTATATCAAGG	1522
Db	1201	GATGTATTCCTCATGATGACACCAACCTTTGAAGAGCAAACTCTTATATCAAGG	1260
QY	1523	ACTGCTGTGCAAGAAATTAATCTTAA	1549
Db	1261	GGGCTGTTGCGAGATTAATCTTAA	1287
RESULT 14			
ABK12990			
ID	ABK12990	standard; DNA; 1287 BP.	
XX	AC	ABK12990;	
XX	AC		
XX	DB	23-APR-2002 (first entry)	
XX	DB		
XX	DB	Synthetic goliath DNA sequence #4.	
XX	DB		
XX	DB	Goliath protein; antiangiogenic; vasotropic; gene therapy;	
XX	DB	dosage form; angiogenesis; neurogenesis; tumour; vascularisation;	

KW cancer; ischemia; neuroblastoma; neurodegenerative disease;
KM unrecovered nerve trauma; ds.
XX Synthetic.
XX
XX W0200193681-A1.
XX
XX
PD 13-DEC-2001.
PN
PP 01-UTN-2001; 2001MO-US18000.
PR 02-UTN-2000; 2000US-0586398.
RA (REGC) UNIV CALIFORNIA.
RX Harland R, Baker JC;
RR WPI, 2002-147637/19.
TX
TX
TX New compositions comprising gollioth proteins, useful for modulating angiogenesis or neurogenesis in mammals e.g. for preventing or treating undesirable vascularrisation of a tumour, ischaemia or neurodegenerative disease -
PT
PS Disclosure, Page 41; 45pp; English.

The present invention relates to a new pharmaceutical composition that comprises a gollioth polypeptide in dosage form. The gollioth polypeptide has a sequence identity of at least 75% to the protein sequences (AAU4918-AAU4921) fully defined in the specification. The composition is useful for modulating angiogenesis or neurogenesis in mammals, particularly in humans or mice. Specifically, the composition is useful for the prophylactic and/or therapeutic treatment of excess angiogenesis e.g. undesirable vascularrisation of a tumour or insufficient angiogenesis e.g. ischaemia, excess neurogenesis e.g. neuroblastoma or insufficient neurogenesis e.g. neurodegenerative disease or unrecovered nerve trauma. The present nucleic acid sequence represents synthetic DNA sequence #4 that is one of several artificial gollioth nucleic acids (ABK12987-ABK12992) used in the invention for modulating angiogenesis or neurogenesis.

Sequence 1287 BP; 312 A; 321 C; 356 G; 298 T; 0 other;

	Query Match	58.8%	Score 1042.8	DB 24	Length 1287
	Best Local Similarity	88.1%	Pred. No. 5.3e-280		
	Matches 1134	Conservative 0	Mismatches 153	Indels 0	Gaps 0
OY	263 ATGGGCGCCGCCCTGTGAGGCCTGTCCTGCGCGGATGCTGCGGCTTTTCAAGATTG	322			
Db	1 ATGGGCGCGCCGCCCTGTGAGGCCTGTCCTGCGCGGATGCTGCGGCTTTTCAAGATTG	60			
OY	323 CTGGGATGAGTGCTTCTCGCTGCGCCCTGAGTCCGACAGCACCGGTTCCCGGGGGCTGA	382			
Db	61 CTGGCTTAGTGCTTCTTCTGCTCTGAGTCCACAGCCCGGTTCCCGGAGCGGA	120			
OY	383 GCAGTGTGACCGCGTAACCTCAACGTCCTGCGCGGTTTCCGACACGGAGTGAACCT	442			
Db	121 GCCGTGTGACCGCGTAACCTCAACGTCCTGCGCGGTTTCCGACACGGAGTGAACCG	180			
OY	443 AGCGTGTGAGCGTGAACGACAGAGAGGCGGTGAAGGACAGACTCGCGCGCTGAGACTGTG	502			
Db	181 ACTGTGTGAGAGCTGAGACACAGAGGCGGTGAAGGACAGACTCGCGCGCTGAGACTGTG	240			
OY	503 GCGTGGGCTCTGTTACCGCCGACCGGCTCCGGGGCGCTTAAGCTGTAAACCGACAAG	562			
Db	241 TCCGGGGCTCTGTTACCGCCGACCGGCTCCGGGGCGCTTAAGCTGTAAACCGACAACC	300			
OY	563 AAATTCAAGGTGACCGAGTTTGGGAGAAGCAAGCTGCAGTCTTGTGGTTGGCCCTCATC	622			
Db	301 AAATTCAAGGTGACCGAGTTTGGGAGAAGCAAGCTGCAGTCTTGTGGTTGGCCCTCATC	360			
OY	623 CAACCGCGCGGCGCTGCACTTGGCAGCAAGATCAATGATGCTTAAGAGATGGGG	682			

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: October 6, 2003, 16:32:57 ; Search time 461 Seconds

9816.120 Million cell updates/sec

Title: US-09-854-300-7

Sequence: 1 agctggagctccaccgcggt.....ataggaagtttcctctcag 1774

Scoring table: IDENTITY_NUC

Searched: 1708419 seqs, 1275431651 residues

Total number of hits satisfying chosen parameters: 3416838

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/prodata/2/pub/pbma/US07_PUBCOMB.seq.*
- 2: /cgn2_6/prodata/2/pub/pbma/US06_NEW_PUB.seq.*
- 3: /cgn2_6/prodata/2/pub/pbma/US06_NEW_PUB.seq.*
- 4: /cgn2_6/prodata/2/pub/pbma/US06_PUBCOMB.seq.*
- 5: /cgn2_6/prodata/2/pub/pbma/US07_NEW_PUB.seq.*
- 6: /cgn2_6/prodata/2/pub/pbma/PC05_PUBCOMB.seq.*
- 7: /cgn2_6/prodata/2/pub/pbma/US08_NEW_PUB.seq.*
- 8: /cgn2_6/prodata/2/pub/pbma/US08_PUBCOMB.seq.*
- 9: /cgn2_6/prodata/2/pub/pbma/US09_PUBCOMB.seq.*
- 10: /cgn2_6/prodata/2/pub/pbma/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/prodata/2/pub/pbma/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/prodata/2/pub/pbma/US09_NEW_PUB.seq.*
- 13: /cgn2_6/prodata/2/pub/pbma/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/prodata/2/pub/pbma/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/prodata/2/pub/pbma/US10_NEW_PUB.seq.*
- 16: /cgn2_6/prodata/2/pub/pbma/US60_NEW_PUB.seq.*
- 17: /cgn2_6/prodata/2/pub/pbma/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query No.	Score	Match	Length	DB	ID	Description
1	1178-8	66.4	2306	10	US-09-764-664-63		Sequence 63, April
2	1157-6	65.0	1250	10	US-09-764-664-519		Sequence 519, April
3	352-6	21.3	643	10	US-09-878-178-1815		Sequence 1815, April
4	373-6	21.0	643	13	US-10-046-935-1815		Sequence 1815, April
5	372-6	21.0	643	14	US-10-104-802-1815		Sequence 1815, April
6	367-4	20.7	509	14	US-10-066-643-1696		Sequence 1696, April
7	320-6	18.1	1188	12	US-09-814-353-21198		Sequence 21198, April
8	310	17.5	435	10	US-09-920-1002-1176		Sequence 1176, April
9	310	17.5	435	12	US-10-039-328-1116		Sequence 1176, April
10	310	17.5	435	13	US-10-039-328-1116		Sequence 1176, April
11	274-6	15.5	387	10	US-09-993-665-5551		Sequence 5551, April
12	224-6	12.7	616	14	US-10-060-936-25		Sequence 25, April
13	214-4	12.1	374	10	US-09-998-665-4662		Sequence 4662, April
14	209	11.8	1200	12	US-10-074-511-3		Sequence 3, April
15	2012	11.3	2063	13	US-09-935-390A-1		Sequence 1, April
16	2012	11.3	392	10	US-09-983-965-5467		Sequence 5467, April

17	187.4	10.6	340	10	US-09-783-559-35523	Sequence 3523, App
18	189.8	10.2	3615	11	US-09-822-849-556	Sequence 596, App
19	149.2	8.4	1322	9	US-09-935-2904-16	Sequence 16, App1
20	147.6	8.3	1253	14	US-10-268-473-6	Sequence 6, App1
21	144.6	8.2	1521	12	US-09-814-353-17086	Sequence 17086, App
22	144.4	8.1	1251	12	US-09-933-767-156	Sequence 156, App
23	144.4	8.1	1251	14	US-10-023-282-156	Sequence 156, App
24	136.4	7.7	442	12	US-09-814-353-4398	Sequence 4398, App
25	136.4	7.7	442	12	US-09-814-353-10702	Sequence 10702, App
26	136.6	7.5	619	10	US-09-764-864-485	Sequence 485, App
27	132	7.4	508	11	US-09-918-995-21070	Sequence 21070, App
28	117.2	6.6	453	11	US-09-864-761-11364	Sequence 11364, App
29	112.8	6.4	826	9	US-09-822-649A-90	Sequence 90, App1
30	112.4	6.3	435	10	US-09-983-965-5079	Sequence 5079, App
31	105.6	6.0	382	9	US-09-825-343-1238	Sequence 1238, App
32	105.2	5.9	253	9	US-09-864-761-27944	Sequence 27944, App
33	102.8	5.8	1074	12	US-09-971-392-211	Sequence 211, App
34	90.4	5.1	2330	11	US-09-822-846-240	Sequence 240, App
35	86.4	4.9	922	9	US-09-800-129-23	Sequence 23, App1
36	80	4.5	1247	14	US-10-106-698-2093	Sequence 2093, App
37	79.4	4.5	549	11	US-09-764-861-126	Sequence 126, App
38	79.2	4.5	455	14	US-10-106-698-1892	Sequence 1892, App
39	78.8	4.5	570	10	US-09-734-011A-7	Sequence 7, App1
40	78.6	4.4	439	9	US-09-926-299-484	Sequence 484, App
41	78.6	4.4	439	11	US-09-925-129-484	Sequence 484, App
42	78.4	4.4	583	12	US-10-160-162-74	Sequence 74, App
43	78.4	4.4	1701	12	US-09-930-213-312	Sequence 312, App
44	78.4	4.4	2346	12	US-10-011-200-2	Sequence 2, App1
45	78	4.4	726	11	US-09-966-480-36	Sequence 36, App1

ALIGNMENTS

```

RESULT 1
US-09-764-864-63
: Sequence 63, Application US/09764864
: Patent No. US20020132753A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: P7223
: CURRENT APPLICATION NUMBER: US/09/764,864
: CURRENT FILING DATE: 2001-01-17
: Prior application data removed - consult PAM or file wrapper
: NUMBER OF SEQ ID NOS: 1792
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO. 63
:
: LENGTH: 2306
:
: TYPE: DNA
: ORGANISM: Homo sapiens
:
: US-09-764-864-63

```

Query Match 66.4%; Score 1178.8; DB 10; Length 2306;

Matches 1217; Conservative 0; Mismatches 8; Indels 4; Gaps 3

Qy	607	TTGGTGGCCCTCATTCMAACGCGCGGCGGCGCTGACCTTCGCAACAAAGTCATCTGCG	666
Db	72	TTGGTGGCCCTCATTCMAACGCGCGGCGCTGACCTTCGCAACAAAGTCATCTGCG	1341
Qy	667	TTTATGAGAGATGGGCGCTGTGAGACCGTCACTTTAACTTCCCGGAGACCGCGAATGAGGT	722
Db	132	TTTATGAGAGAGGCGGCTGTGAGACCGTCACTTTAACTTCCCGGAGACCGCGAATGAGGT	1919
Qy	727	CATCCCATGTCATACC CGGSGTGAAGTGA CATTGTGGATCATGATCGGCAATCTGA	786
Db	132	CATCCCATGTCATACC CGGSGTGAAGTGA CATTGTGTGCATCATGATCGGCAATCTGA	2515

QY 787 AGGCAAAAATTCGCAATCTATTCAAGAGGCAATACAGATGCAATAGTAGT 846
DB 252 AGGCAAAAATTCGCAATCTATTCAAGAGGCAATACAGATGCAATAGTAGT 311
QY 847 AGGCAAAAATTCGCAATCTATTCAAGAGGCAATACAGATGCAATAGTAGT 905
DB 312 AGGCAAAAATTCGCAATCTATTCAAGAGGCAATACAGATGCAATAGTAGT 371
QY 906 TTTTATTTATTTAGGCGGCACTGTGGGCTATTTATCTTTATCTGCTCGAAGGCTAC 965
DB 372 TTTTATTTATTTAGGCGGCACTGTGGGCTATTTATCTTTATCTGCTCGAAGGCTAC 431
QY 966 GGAATGCAAGGCTTCAAGCAAGAGCAAGAGCAATTAAGGCAAGTCAAAAGCTA 1025
DB 432 GGAATGCAAGGCTTCAAGCAAGAGCAAGAGCAATTAAGGCAAGTCAAAAGCTA 491
QY 1026 TTGGAAGGCTTCAAGCAAGAGCAAGAGCAATTAAGGCAAGTCAAAAGCTA 1085
DB 492 TTGGAAGGCTTCAAGCAAGAGCAAGAGCAATTAAGGCAAGTCAAAAGCTA 551
QY 1086 ATAGTGTGCTGTGCAATGATTAATTAACCAATGATTTGTAGCATCTTAAGT 1145
DB 552 ATAGTGTGCTGTGCAATGATTAATTAACCAATGATTTGTAGCATCTTAAGT 611
QY 1146 GCAACCAATTTTCCATTAAGCAATGCTGCAATGCTGCTGCTGCTGCTGCTGCT 1205
DB 612 GCAACCAATTTTCCATTAAGCAATGCTGCAATGCTGCTGCTGCTGCTGCTGCT 671
QY 1206 CCATGCAATTTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1265
DB 672 CCATGCAATTTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 731
QY 1266 CAGTGTCTTTACCAAGTCCCTGATTCATCAATGATTAATGATGCTGCTGCTGCTGCT 1325
DB 732 CAGTGTCTTTACCAAGTCCCTGATTCATCAATGATTAATGATGCTGCTGCTGCTGCT 791
QY 1326 AGGATATTCGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGG 1385
DB 792 AGGATATTCGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGG 851
QY 1386 CGCCTCTGAGAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAG 1445
DB 852 CGCCTCTGAGAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAG 911
QY 1446 CAAATTCGTGCGAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGG 1505
DB 912 CAAATTCGTGCGAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGG 971
QY 1506 AAACCTCCTAATCAAGAGCTGCTGTGAGAGAAATTAATCTTAATAATCTGTAAATGA 1565
DB 972 AAACCTCCTAATCAAGAGCTGCTGTGAGAGAAATTAATCTTAATAATCTGTAAATGA 1031
QY 1566 AAACCTGCAACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1625
DB 1032 AAACCTGCAACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1091
QY 1626 TGGATTAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1685
DB 1092 TGGATTAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1151
QY 1686 ATAGTGTGCTGTGCAATGATTAATTAACCAATGATTTGTAGCATCTTAAGT 1745
DB 1152 ATAGTGTGCTGTGCAATGATTAATTAACCAATGATTTGTAGCATCTTAAGT 1208
QY 1746 ATATTTTTCATAGGCAAGTTCCTCTGAG 1774
DB 1209 ATATTTTTCATAGGCAAGTTCCTCTGAG 1237

RESULT 2
US-09-764-864-519
Sequence 519, Application US/09764864
Patent No. US20020132753A1

GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: Ptz23
CURRENT APPLICATION NUMBER: US/09/764,864
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PAM or file wrapper
NUMBER OF SEQ ID NOS: 1792
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 519
LENGTH: 1250
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (540)
OTHER INFORMATION: n equals a,t,g, or c
LOCATION: (1242)
OTHER INFORMATION: n equals a,t,g, or c
US-09-764-864-519
Query Match 65.3%; Score 1157.6; DB 10; Length 1250;
Best Local Similarity 98.8%; Pred. No. 1.3e-308;
Matches 1208; Conservative 0; Mismatches 10; Indels 5; Gaps 4;
QY 547 CTGTACCGCAAGCAATTTCAAGTCCCAAGGTTGGGAAAGCAAGGCAAGTCTC 606
DB 18 CTGTACCGCAAGCAATTTCAAGTCCCAAGGTTGGGAAAGCAAGGCAAGTCTC 77
QY 607 TTGTGTGGCCCTTATTCACAGCGCGCGGGCTGCACTTCCCAAGCAAGTCCATGCG 666
DB 78 TTGTGTGGCCCTTATTCACAGCGCGCGGGCTGCACTTCCCAAGCAAGTCCATGCG 137
QY 667 TTATGAGATGGGCGCTGAGAGCCCTCATCTTAACTTCCCGGCAAGGCAAGTGGT 726
DB 138 TTATGAGAGAGGGGGCTGAGAGCCCTCATCTTAACTTCCCGGCAAGGCAAGTGGT 197
QY 727 CATCCCAATGTCTACCGCGGGCTGAGAGCAATGTTGCAATGATCGGCAATGGA 786
DB 196 CATCCCAATGTCTACCGCGGGCTGAGAGCAATGTTGCAATGATCGGCAATGGA 257
QY 787 AGGCAAAAATTCGCAATCTATTCAAGAGGCAATCAAGTGAATATGCTATTAAGT 846
DB 256 AGGCAAAAATTCGCAATCTATTCAAGAGGCAATCAAGTGAATATGCTATTAAGT 317
QY 847 AGGCAAAAATTCGCAATCTATTCAAGAGGCAATCAAGTGAATATGCTATTAAGT 905
DB 318 AGGCAAAAATTCGCAATCTATTCAAGAGGCAATCAAGTGAATATGCTATTAAGT 377
QY 906 TTTTATTTATTTAGGCGGCACTGTGGGCTATTTATCTTTATCTGCTCGAAGGCTAC 965
DB 378 TTTTATTTATTTAGGCGGCACTGTGGGCTATTTATCTTTATCTGCTCGAAGGCTAC 437
QY 966 GGAATGCAAGGCTTCAAGCAAGAGCAAGAGCAATTAAGGCAAGTCAAAAGCTA 1025
DB 438 GGAATGCAAGGCTTCAAGCAAGAGCAAGAGCAATTAAGGCAAGTCAAAAGCTA 497
QY 1026 TTGGAAGGCTTCAAGCAAGAGCAAGAGCAATTAAGGCAAGTCAAAAGCTA 1084
DB 498 TTGGAAGGCTTCAAGCAAGAGCAAGAGCAATTAAGGCAAGTCAAAAGCTA 557
QY 1085 GGAATGCTGCTGTGCAATGATTAATTAACCAATGATTTGTAGCATCTTAAGT 1144
DB 558 GGAATGCTGCTGTGCAATGATTAATTAACCAATGATTTGTAGCATCTTAAGT 617
QY 1145 TGAACCAATTTTCCATTAAGCAAGTGTGCAAGGCAAGGCTTCAAAAGCAAGCTTCG 1204
DB 618 TGAACCAATTTTCCATTAAGCAAGTGTGCAAGGCAAGGCTTCAAAAGCAAGCTTCG 677
QY 1205 CCATGCAATTTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1264
DB 678 CCATGCAATTTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 737

1723 0C1G0IAHACUIGWABUCMAI I CIGIBCBAGIBGAIGINAI CCICATGINWAAHACLLC 1480

1609 AGITTCIAITAAAIIGSAIAAHTTAAATAAAATAAGAGIGATACIGAHAGIGCICAG 1600

Db 241 AGTTCTATTAAATGGAATAATTATATAAATAAGAGTACTGAAAGTCCTCAG 300


```

; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21198
; LENGTH: 1188
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-353-21198

```

```

Query Match      18.1%; Score 320.6; DB 12; Length 1188;
Best Local Similarity 97.8%; Pred. No. 1.5e-77;
Matches 357; Conservative 0; Mismatches 4; Indels 4; Gaps 3;

```

```

QY 1410 CAACAATGAAAGCTACAGCTGTGTAACGAAAGCAATTCGTGCGAGTGAATTA 1469
DB 1 CAACAATGAAAGCTACAGCTGTGTAACGAAAGCAATTCGTGCGAGTGAATTA 60
QY 1470 TTCCTCATGTGACCAACCACTTTGAAAGAGCAACTCTTAATCAAGAGCTGCTG 1529
DB 61 TTCTCATGTG-GACCAACCACTTTGAAAGAGCAACTCTTAATCAAGAGCTGCTG 119
QY 1530 TTCGAAATTAATCTTAATCTGTTAAATGTAAGAAAAGTTGACCATTAATTAACA 1589
DB 120 TTCGAAATTAATCTTAATCTGTTAAATGTAAGAAAAGTTGACCATTAATTAACA 179
QY 1590 GAACTGCAATCAGAGGCTAGTCTTATTAATTAATTAATTAATTAATTAATTAATTAAG 1649
DB 180 GAACTGCAATCAGAGGCTAGTCTTATTAATTAATTAATTAATTAATTAATTAATTAAG 239
QY 1650 TGATACCTGAAAGTGTCTGATGACTTAATTAATTAATTAATTAATTAATTAATTAAT 1709
DB 240 TGATACCTGAAAGTGTCTGATGACTTAATTAATTAATTAATTAATTAATTAATTAAT 297
QY 1710 TTAACCTGTTAATCTTTTTCACCAACCTTAATTAATTAATTAATTAATTAATTAATTAAT 1769
DB 298 TTAACCTGTTAATCTTTTTCACCAACCTTAATTAATTAATTAATTAATTAATTAATTAAT 356
QY 1770 CTCAG 1774
DB 357 CTCAG 361

```

```

RESULT 8
US-09-920-300A-1176/c
; Sequence 1176, Application US/09920300A
; Patent No. US20020136728A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon B.
; APPLICANT: Mesgher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547
; CURRENT APPLICATION NUMBER: US/09/920,300A
; NUMBER OF SEQ ID NOS: 1789
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1176
; LENGTH: 435
; TYPE: DNA
; ORGANISM: Homo sapiens

```

```

US-09-920-300A-1176

```

```

Query Match      17.5%; Score 310; DB 10; Length 435;
Best Local Similarity 97.4%; Pred. No. 6.2e-75;
Matches 336; Conservative 0; Mismatches 6; Indels 3; Gaps 2;

```

```

QY 1430 CTGGTAAACCAAGAGCAATTCGTGCGAGTGAATTAATTCCTCATGTGACCAACCA 1489
DB 435 CTGGTAAACCAAGAGCAATTCGTGCGAGTGAATTAATTCCTCATGTGACCAACCA 376
QY 1490 ACCTTGAAAGAGCAAGCAACTCTTAATCAAGAGCTGCTGTTGAAGAAATTAATCTTAA 1549
DB 375 ACCTTGAAAGAGCAAGCAACTCTTAATCAAGAGCTGCTGTTGAAGAAATTAATCTTAA 316
QY 1550 AATCTGTAAATTAAGAAATCTGAACCTTAATTAATTAATTAATTAATTAATTAATTAAT 1609
DB 315 AATCTGTAAATTAAGAAATCTGAACCTTAATTAATTAATTAATTAATTAATTAATTAAT 256
QY 1610 GTTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1669
DB 255 GTTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 196
QY 1670 TGACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1729
DB 195 TGACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 138
QY 1730 CACCAACTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1774
DB 137 CA-CACCAACTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 94

```

```

RESULT 9
US-10-099-926-1176/c
; Sequence 1176, Application US/10099926
; Publication No. US2003016064A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon B.
; APPLICANT: Mesgher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; APPLICANT: Jiang, Yugu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547C2
; CURRENT APPLICATION NUMBER: US/10/099,926
; NUMBER OF SEQ ID NOS: 1992
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1176
; LENGTH: 435
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-099-926-1176

```

```

Query Match      17.5%; Score 310; DB 12; Length 435;
Best Local Similarity 97.4%; Pred. No. 6.2e-75;
Matches 336; Conservative 0; Mismatches 6; Indels 3; Gaps 2;

```

```

QY 1430 CTGGTAAACCAAGAGCAATTCGTGCGAGTGAATTAATTCCTCATGTGACCAACCA 1489
DB 435 CTGGTAAACCAAGAGCAATTCGTGCGAGTGAATTAATTCCTCATGTGACCAACCA 376
QY 1490 ACCTTGAAAGAGCAAGCAACTCTTAATCAAGAGCTGCTGTTGAAGAAATTAATCTTAA 1549
DB 375 ACCTTGAAAGAGCAAGCAACTCTTAATCAAGAGCTGCTGTTGAAGAAATTAATCTTAA 316
QY 1550 AATCTGTAAATTAAGAAATCTGAACCTTAATTAATTAATTAATTAATTAATTAATTAAT 1609
DB 315 AATCTGTAAATTAAGAAATCTGAACCTTAATTAATTAATTAATTAATTAATTAATTAAT 256
QY 1610 GTTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1669
DB 255 GTTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 196

```

QY 1670 TGAATAATATGCTATAGTAAATGAGCTTAAATATTAACCTGTTAACTTTTTC 1129
DB 195 TGAATAATATGCTATAGTAAATGAGCTTAAATATTAACCTGTTAACTTTTTC 138
QY 1730 CACCAACTCATTAATATATTTTTCATAGGCAAGTTCCTCTGAG 1774
DB 137 CA-CAACTCATTAATATATTTTTCATAGGCAAGTTCCTCTGAG 94

RESULT 10

US-10-033-528-1176/c
Sequence 1176, Application US/10033528
Publication No. US20020131971A1
GENERAL INFORMATION:
APPLICANT: King, Gordon E.
APPLICANT: Mesgher, Madeleine Joy
APPLICANT: Xu, Jiangchun
APPLICANT: Secretist, Heather
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
FILE REFERENCE: 210121.547C1
CURRENT APPLICATION NUMBER: US/10/033.528
CURRENT FILING DATE: 2001-12-26
NUMBER OF SEQ ID NOS: 1896
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1176
LENGTH: 435
TYPE: DNA
ORGANISM: Homo sapiens
US-10-033-528-1176

Query Match 17.5%; Score 310; DB 13; Length 435;
Best Local Similarity 97.4%; Pred. No. 6.2e-75;
Matches 336; Conservative 0; Mismatches 6; Indels 3; Gaps 2;

QY 1430 CTGTAACCATGAGCAAAATTCGTGGCAGTGGATTTTCTCATGTMGACCAACCA 1489
DB 435 CTGGTAACCATGAGCAAAATTCGTGGCAGTGGATTTTCTCATGTMGACCAACCA 376
QY 1490 ACCTTGAGAGAGCAAACTCTCATCAAGAGCTGTTCCAGAAATTAATCTTAA 1549
DB 375 ACCTTTGAGAGAGCAAACTCTCATCAAGAGCTGTTCCAGAAATTAATCTTAA 316
QY 1550 AATCTGTGTAATAGAGAACTTGAACATTAAGTAATAGAACTGCAATCAGGAGCTTA 1609
DB 345 AATCTGTGTAATAGAGAACTTGAACATTAAGTAATAGAACTGCAATCAGGAGCTTA 256
QY 1610 GTTCTATTAATATGAGTAATATTAATTAATTAAGAGTGAATGCAAGTCTCAGA 1669
DB 255 GTTCTATTAATATGAGTAATATTAATTAATTAAGAGTGAATGCAAGTCTCAGA 196
QY 1670 TGAATAATATGCTATAGTAAATGAGCTTAAATATTAACCTGTTAACTTTTTC 1729
DB 195 TGAATAATATGCTATAGTAAATGAGCTTAAATATTAACCTGTTAACTTTTTC 138
QY 1730 CACCAACTCATTAATATATTTTTCATAGGCAAGTTCCTCTGAG 1774
DB 137 CA-CAACTCATTAATATATTTTTCATAGGCAAGTTCCTCTGAG 94

RESULT 11

US-09-983-965-5551
Sequence 5551, Application US/09983965
Patent No. US20020137160A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing
APPLICANT: Byatt, John C.
APPLICANT: Machalagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
FILE REFERENCE: 37-21 (10297) C

CURRENT APPLICATION NUMBER: US/09/983,965
CURRENT FILING DATE: 2001-10-26
PRIOR APPLICATION NUMBER: US 09/465,231
PRIOR FILING DATE: 1999-12-15
PRIOR APPLICATION NUMBER: US 60/113,678
PRIOR FILING DATE: 1998-12-17
NUMBER OF SEQ ID NOS: 5912
SEQ ID NO 5551
LENGTH: 387
TYPE: DNA
ORGANISM: Bos taurus
FEATURE:
NAME/KEY: unsure
LOCATION: (347)
OTHER INFORMATION:
OTHER INFORMATION: Clone ID: 50-LIB34-010-Q1-E2-E6
US-09-983-965-5551

Query Match 15.5%; Score 274.6; DB 10; Length 387;
Best Local Similarity 86.8%; Pred. No. 3.4e-65;
Matches 336; Conservative 0; Mismatches 46; Indels 5; Gaps 3;

QY 1386 CGCTCTGAGGAAACGCTGAGTCAACAATGAAGTTCAGCTGTTAACTGTAAG 1445
DB 1 CACCTTGAAGGAAACGCTGAGTCAACAATGAAGTTCAGCTGTTAACTGTAAG 60
QY 1446 CAATTCGTGGCAGTGGATTTATCTCATGTAAGCAACCAACCTTGAAGAGAG 1505
DB 61 CAATTCGTGGCAGTGGATTTATCTCATGTAAGCAACCAACCTTGAAGAGAG 120
QY 1506 AATCTGTGTAATAGAGAACTTGAACATTAAGTAATAGAACTGCAATCAGGAGCTTA 1565
DB 121 AATCTGTGTAATAGAGAACTTGAACATTAAGTAATAGAACTGCAATCAGGAGCTTA 180
QY 1566 AATCTGAAACCTTGAATTAAGAGAACTGCAATCAGGAGCTTAACTTAA 1624
DB 181 AATCTGAAACCTTGAATTAAGAGAACTGCAATCAGGAGCTTAACTTAACTTAA 240
QY 1625 TTGATTAATTAATTAATTAAGAGTTCAGAAAGTCTCAAGTCAATTAATAGC 1684
DB 241 CTGATTAATTAATTAATTAAGAGTTCAGAAAGTCTCAAGTCAATTAATAGC 300
QY 1685 TATGTAATTAATGCTTAAATATTAATTAACCTGTTAACTTTTTCACCAACTCATTT 1744
DB 301 TATGTAATTAATGCTTAAATATTAATTAACCTGTTAACTTTTTCACCAACTCATTT 356
QY 1745 AATATTTTCAATGAGCAAGTTCCTCT 1771
DB 357 AATATTTTCAATGAGCAAGTTCCTCT 383

RESULT 12

US-10-060-036-25
Sequence 25, Application US/10060036
Publication No. US20030073144A1
GENERAL INFORMATION:
APPLICANT: Benson, Darin R.
APPLICANT: Kalos, Michael D.
APPLICANT: Iodes, Michael J.
APPLICANT: Persing, David H.
APPLICANT: Hepler, William T.
APPLICANT: Jiang, Yudi
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER
FILE REFERENCE: 210121.566
CURRENT APPLICATION NUMBER: US/10/060,036
CURRENT FILING DATE: 2002-01-30
NUMBER OF SEQ ID NOS: 4560
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 25
LENGTH: 616
TYPE: DNA
ORGANISM: Homo sapiens

Dy 699 GACCTTCGAGCAAGATCATTTGGCTTATAGAGGTGGCGTCTGAGCCGCATCT 698
| | | | | | | | | | | | | |
Db 353 GACCTTCAGAGCAAGTGCTGTGGGCGCGGAGSAAAGCCTCGGCCCTCGTCTCT 412

```

QY      639  TTAACTTCCCGGGAACCGGCAATGAGTATCCCATGTCTTCAACCCGGGTGAGTAACA 758
D      413  ACAATGAGAGGCGCTACGGAACATCACTTGCATGTCAAGCGGGAACAGGAATA 472
QY      759  TTGTGCAATCATGATGCGCAATGTGAAAGCAAAAATCTGCATCTATTTCAAAG 818
D      473  TAGTGTCATTAATGATTAAGTATCAAAAAGAAATTTGAGAGCTGTGCAAAAG 532
QY      819  GCATACAGTGAATGTCATGATGAGTAAGTGAAGGAAAACATGAGCTTGGGTAATCA 877
D      533  GAATTCAGTAACGATGACCATAGGGGTGCAACCGGCATGTACAGAGATTCATCAACG 592
QY      878  --TATCAATTTTTCGTCTGTGTCCTTTTATTAATTAAGGGGCAACTGAGGCT 935
D      593  GTGAGTCTGTGAGTGTGAGCTTGCCTTATCAATCATATATATCTGTAGCT 652
QY      936  ATTTTATCTTTTATCTGTCTGAAAGCTACGAAATGCAAGAGCTCAAAAGAGAAAGCA 995
D      653  GGCTAATATTTTACTATTAACAGCGTTTCTATACTAGGCTCAGATTGCAAGTCA 712
QY      996  GGCAATTAAGGAGATGCTAAAAAGCTATTGGAAGGCTCAACTAAGCACTGAAG 1055
D      713  GCATAGAAA--AGAACTAAGAAATTAATGAGTATGAGCTTACTTCACTATCTGAAGC 769
QY      1056  AAGGAGCAAGAAATGAGCCCTGATGAGAAATGTTGCTGTGTGCAATTAATGTATA 1115
D      770  ATGAGAAAAGGAAATGATGTATGCTGAATATGTCAGTGTGTAATGAATTTCA 829
QY      1116  AACCAATGATTTGGAGCAATCTTAAGTGAACCAATTTCCATTAAGCACTGTGTG 1175
D      830  AAGTAAGAAATTAATTAAGATTTGCCATGCAAGATTTTCAATGAATGATG 889
QY      1176  ACCCATGCTGTTAAACCAAGACTTGCCTCCATGTGAATGCAATCTCAAGCT 1235
D      890  ACCCATGCTTGTGATCAAGCAATGTCATGTAATGTAATGATGATCAAGGCC 949
QY      1236  TGGGATTTAGGTGATGTTGAAGTGA 1264
D      950  TAGGATTTGGGAGAGCTGGGAGTGA 978

```

```

RESULT 15
US-09-935-390A-1
; Sequence 1, Application US/09935390A
; Patent No. US20020076761A1
; GENERAL INFORMATION:
; APPLICANT: Becbedo, Jaime
;              Quianjin, Hu
;              Garcia, Pablo
;              Williams, Lewis T.
;              Kothakota, Srinivas
; TITLE OF INVENTION: Secreted Human Proteins
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/935,390A
; FILING DATE: 22-Aug-2001
; CLASSIFICATION: <Unknown>
; *PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/998,671
; FILING DATE: 1997-12-11
; ATTORNEY/AGENT INFORMATION:

```

```

; NAME: Jane E. R. Potter
; REGISTRATION NUMBER: 35,332
; REFERENCE/DOCKET NUMBER: 1369.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 923-2718
; TELEFAX: (510) 655-3542
; TREX: <Unknown>
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2063 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-935-390A-1

Query Match      11.8%; Score 209; DB 9; Length 2063;
Best Local Similarity 54.6%; Pred. No. 1,5e-46;
Matches 507; Conservative 0; Mismatches 410; Indels 12; Gaps 4;

QY      339  TGCCTGCGCTGAGTCCGACGACCGCGTTCCCGGGGCGCTGAAGCAGTGGACCGGCT 398
D      180  TGCTGCTGCGCTGAGTCCGACGACCGCGTTCCCGGGGCGCTGAAGTGGTTCG 229
QY      399  ACTCAAGTGTCTGCGCGGCTTCGACACGGAAGTGAACCTGACGTGTGAGAGTGA 458
D      240  CCGT---GGTAAATCATCAGATGATGAGACCGCAACCACTGACGCGTGTGAGAGCT 296
QY      459  GCGAGAGGCGCGTGAAGGCGCAAGACTGCGCTGGAAGCGTGGGCGGCTGTGATAC 518
D      297  CGAGAGTGTGCGCTTCGCGCAAGCTGCCCAAGAGAGGCGCGCAATGCGCTGTGAG 356
QY      519  CGCCCAAGCGCGCGCGCGCTTAAGCGCTGAACCGCAACAGAAATTCAGAGTCCCA 578
D      357  TCCCGTGGCGCGCGCGCGCGAGACCTGAGAGCGCTGCGCGCGCGCAACCGCTTCTTG 416
QY      579  CGGTTTGGGGAAGCAACGTGCAAGTCTCTTGTGGCGCTCATCAACCGCGCGGAGCT 638
D      417  CGAGAGCGCGCGCGCGCGCGCGCGCGCGCTTGTGCTGCTGTGCTGCTGCTGCTGCT 473
QY      639  GCACCTTCGCAACAAATTCATCTGCTTATGAGAGTGGCGCTCTGAGACCTGTATCT 698
D      474  GCACCTTCGCAACAAATTCATCTGCTTATGAGAGTGGCGCTCTGAGACCTGTATCT 533
QY      699  TTAACTTCCCGGGAACCGGCAATGAGTATCCCATGTCTCAACCGGAGTGAACA 758
D      534  ACAATGAGAGGCGCTACGGAACATCACTTGCATGTCAAGCGGGAACAGGAATA 593
QY      759  TTGTGCAATCATGATGCGCAATGTGAAAGCAAAAATCTGCATCTATTTCAAAG 818
D      594  TAGTGTCATTAATGATTAAGTATCAAAAAGAAATTTGAGAGCTGTGCAAAAG 653
QY      819  GCATACAGTGAATGTCATGATGAGTAAGTGAAGGAAAACATGAGCTTGGGTAATCA 877
D      654  GAATTCAGTAACGATGACCATAGGGGTGCAACCGGCATGTACAGAGATTCATCAACG 713
QY      878  --TATCAATTTTTCGTCTGTGTCCTTTTATTAATTAAGGGGCAACTGAGGCT 935
D      714  GTGAGTCTGTGAGTGTGAGCTTGCCTTATCAATCATATATATCTGTAGCT 773
QY      936  ATTTTATCTTTTATCTGTCTGAAAGCTAAGCAAGCAAGCTCAAGAGCAAGAAAGCA 995
D      774  GGCTAATATTTTACTATTAACAGGTTTCCATTAATGAGTGTCAATGGAAGTCA 833
QY      996  GGCAATTAAGGAGATGCTAAAAAGCTATTGGAAGGCTCAACTAAGCACTGAAG 1055
D      834  GCATAGAAA--AGAACTAAGAAATTAATGAGTATGAGCTTCTCACTCACTGAAGC 890
QY      1056  AAGGAGCAAGAAATGAGCCCTGATGAGAAATGTTGCTGTGATGAATGTGTATA 1115
D      891  ATGAGAAAAGGAAATGATGTATGATGCTGAATATGTCAGTGTGTAATGAATTTCA 950

```


569 ACGTCCCAAGGTTTGGGGAAGCACCGGCAAGTCTCTGGTTGGCCCTCATCCAGGC 628
 514 ACGAGGCCA-----AACAGGAGATCTATGGCTGCGCTCATCAACAC-- 556
 629 GCGGGGGGCTGCACTTCGACAGACAAATCCATCTGGCTTATGAGAAATGGCGTCTGA 688
 557 -GTGAGGCTGTACTTTTACATTAATCAAGTGGGACAGAGAGGAGCAATATGG 615
 689 GCGGTCACTTTAACTCCCGGAGACCGGCAATGAGGTCAATCCCATCTCCACCGGCT 748
 616 GTGATCATCTTCAACATCTTCAAGTACGGGCGATAGTATTTCCCATCTCCACAGGG 675
 749 GCAATGACATTTTGCATTCATCATGATCGGCAATCTGAAAGGCAAAATTTCTGCACT 808
 676 ACGAAATATAGTCGCGGTGATGATAGCAACCTGAAAGGCAATGAAATTTGCACTG 735
 809 ATTCAAGAGGCTTACAGATGACATTCGTCATGAGTGGGAGAGAGAGAGAGAGAGAG 868
 736 ATTCAAGAGGAGTCTTGTGACATCTCATCTTGAATGGGAGAGAGAGAGAGAGAGAG 795
 869 GTGATTCATCTTCAATTTTCTGTTCTGTTCTGTTCTTCTTCTTCTTCTTCTTCTTCT 928
 796 GTGACCATTTACATCATGATCTTATTTACCTTCTGCGCCACCAATTTGCTTCTTAC 855
 929 GTGGCTATTTTATCTTTATCTGCTGAGGCTACGGAATGCAAGGCTCAAGAGAG 988
 856 TTAGATGCGTCTGAGACCTTACCTTAGAGTG-----CCAAATCTTCTTCAAGAG 906
 989 AAGCAGAGCAATTTAAGGCAATGCTTAAAGGCTTTTGAAGGCTTCACTACAGACA 1048
 907 AGCGAGGCTCAAAATTAAGGCAATGCTTGAAGGCTTGAAGGCTTCACTACAGAGTT 966
 1049 CTGAACAAGGAGACAGGAAATTTGCGCTGAGGATGATGCTGCTGCTGCTGCTGCTG 1108
 967 CTGAAGAGGAGGATGAGGATTTAAGCTTAAGGAGACATGATGCTGCTGCTGCTGCTG 1026
 1109 TGTGTAACCAATGATTTGCTGAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGG 1168
 1027 AATACCAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAG 1086
 1169 TGTGTAACCAATGATTTGCTGAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGG 1228
 1087 TGTGTAACCAATGATTTGCTGAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGG 1146
 1229 AAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGA 1287
 1147 AATACCAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAG 1206
 1288 AATACCAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAG 1313
 1207 AATACCAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAG 1232

RESULT 2
 US-08-786-606-6
 ; Sequence 6, Application US/08786606
 ; Patent No. 5861495
 ; GENERAL INFORMATION:
 ; APPLICANT: Hillman, Jennifer L.
 ; APPLICANT: Au-Young, Janice
 ; APPLICANT: Coleman, Roger
 ; APPLICANT: Goli, Surya K.
 ; TITLE OF INVENTION: NOVEL HUMAN ZINC-BINDING
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: Fast-Seq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/786, 606
 FILING DATE:
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy RJ
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0173 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 TELEX:
 INFORMATION FOR SEQ. ID NO. 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1253 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-786-606-6

Query Match 8.3%; Score 147.6; DB 2; Length 1253;
 Best Local Similarity - 55.9%; Pred. No. 5,2e-31;
 Matches 301; Conservative 0; Mismatches 234; Indels 3; Gaps 1;
 720 ATGAGTATCCCATCCATGCTTCCCGGAGTGAACATTTGTCATGATGATGAGCA 779
 242 AGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 301
 780 ATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 839
 302 AATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 361
 840 TGAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 896
 362 TACGCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 421
 897 CTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 956
 422 CAATTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 481
 957 GAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1016
 482 AGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 541
 1017 AAAAGCTATTTGAGGCTTCACTACGACATGAAACAGAGAGCAAGAAATTTGGCC 1076
 542 AGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 601
 1077 CTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1136
 602 CAGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 661
 1137 TCTTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1196
 662 TTTCTCCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 721
 1197 AATCTGCCCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1254
 722 GTACCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 779

RESULT 3
 US-09-016-434-5
 ; Sequence 5, Application US/09016434
 ; Patent No. 6500938
 ; GENERAL INFORMATION:

APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSER: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 845-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1253 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: THERPLIB01
CLONE: 010773
US-09-016-434-5

Query Match 8.3%; Score 147.6; DB 4; Length 1253;

Best Local Similarity 55.9%; Pred. No. 5.2e-31;

Matches 301; Conservative 0; Mismatches 234; Indels 3; Gaps 1;

QY 720 ATGAGTCATCCCATGTCACCCGGGTCAGTACATTTGCAATCATGATGAGCA 779
DB 242 AGGAGCCAGTACATGATCATCCAGGCACTGGAGATATTATGCTGTGATGATACG 301
QY 780 ATCTGAAGGCACAAAATCTGCATCTATTCAAAGGAGATCAAGTACAGATGCA 839
DB 302 AATTGAGGGGTAGAGATTTTGAGTTATCTGAGAGAAAATCTGTGACAAAGACAA 361
QY 840 TGAAGTAGGAAAAAACAATGAGCCCTTGGG--TGATCATCTATTCATTTTTCGTT 896
DB 362 TAGCTGTGAGACTGCAAGACCAAGAACTTACGCGTGGCTCTAGCTCTCGGT 421
QY 897 CTGTGCTTTTATTTATTTACGGGCACTGTTGGGCTATTTATCTTTATTTTGTGCTC 956
DB 422 CAATATCTTTATTTTGTGATGATTTTCTTCAGCATGCTCATTTCTTCATCTC 481
QY 957 GAAAGCTACGGAATGACAGCTCAAAAGCGAGAGACAGCAATTTAAAGCAGATGCTA 1016
DB 482 AGAAGATCAGGTACCAAAATGACGACAGGAAACAGGCTGCTGAGATGAGGCA 541
QY 1017 AAAAGCTATTGGAAGGCTTCACTACGCACTGAAAACAAGAGCAAGAAATTTGACC 1076
DB 542 AGAAGGCATCAGTAATGACACACGAGACAGTAAGAAGGAGGAAAGAAATGAGCC 601
QY 1077 CTGATGAGATAGTGTGCTGTGTGCTGATGATGATGATTAACCAATGATTTGTAGCA 1136

DB 602 CAGACTTGATCATTTGAGAGTCTGACATAGAGAGCATTAAGAGATGATGTCGGA 661
QY 1137 TCTTAAGTGAACCAATTTTCCATAGACAGTGTACCCATGCTGTAAACACA 1186
DB 662 TTCTCCCTGAGACATGTTTCCAGAAATCTGCGTGAATCCGCGTTAGGAACAT 721
QY 1197 AGACTTCCCATGAGCAATGTGACATCTCAAGCTTTGGGAATGAGTGTATGT 1254
DB 722 GTACCTGTCTATGTGCAACTTATATTAAGAGCCCTGGAATTTGTCGCAATTT 779

RESULT 4

US-09-205-258-156

Sequence 156, Application US/09205258

Patent No. 6525174

GENERAL INFORMATION:

APPLICANT: Young et al.

TITLE OF INVENTION: 207 Human Secreted Proteins

FILE REFERENCE: P2007P1

CURRENT APPLICATION NUMBER: US/09/205,258

EARLIER FILING DATE: 1998-12-04

EARLIER APPLICATION NUMBER: PCT/US98/11422

EARLIER FILING DATE: 1998-06-04

EARLIER APPLICATION NUMBER: 60/048,885

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,375

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,881

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,880

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,896

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,020

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,876

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,895

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,884

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,894

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,971

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,964

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,882

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,899

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,893

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,900

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,901

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,892

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,915

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,019

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,970

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,972

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,916

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,373

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,875

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,878
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/070,923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 156
LENGTH: 1251
TYPE: DNA
ORGANISM: Homo sapiens
US-09-205-258-156

Query Match 8.1%; Score 144.4; DB 4; Length 1251;
Best Local Similarity 55.6%; Pred. No. 4e-30; Mismatches 236; Indels 3; Gaps 1;
Matches 299; Conservative 0;
Query 720 ATGAGTGCATCCCATGTCACCCGGGTCAGTAGACATGGTTCATCATGATGCGCA 779
Db 189 AGGAGCCAGTACCACTGACTCCAGGCACTGAGCATATATGCTGATGATACG 248
Query 780 ATCTGAAGGACAAATTCGCAATCTTAAGAAGGATCAAGTGAACAGTCA 839
Db 249 AATTGAGGGTAAAGATATTTGATATCTGAGAAAGATCTCTCAAAAGACA 308
Query 840 TGAAGTAGGAAAAACATGGCCCTTGG--GTGATCACTATTCATTTTTCGTT 896
Db 309 TAGCTGTGAACTCGAATGCAACCGAAGAACTTCAGCCGCTCTGATGTCGT 368
Query 897 CTGTGCTTTTATTTATTTACGCGGCACTGAGGCTATTTATCTTTATTCGTC 956
Db 369 CATATCTTTATTTGTTGATGATATTTCTTCAGCATGCTCATNTTCAATC 428
Query 957 GAAGCTACGGAATGCAAGACTCAAGCGAGAGCAAGGCAATTAAGCGCATGCTA 1016
Db 429 AGAAGATCAGGTACAAATGCAACGCGCAAGGAAACCAAGGCTGCGAGATGCGCA 488
Query 1017 AAAAGCTATGAAAGGCTTCAACTACGCACTGMAAACAGAGACAGAGAAATTTGCC 1076
Db 489 AGAAGCATCATGTAATTTGACACACAGACAGTAAGAAAGAGGTGACAGAGAACTGACC 548
Query 1077 CTGATGAGATAGTGTGCTGTGCGCATGTAATTTGTAACCAATGATTTGTTACGA 1136
Db 549 CAGACTTTGATCATGCGAGTCTGCAATGAGAGCTATGACAGATGATGTGTCGAA 608
Query 1137 TCTTAAGTGCACCATATTTTTCATTAAGACATGTTTCAACCATGATGCTGTTAAACACA 1196
Db 609 TTCTCCCTGCAAGCATGTTTTCACAAATCTGCGGATCCCTGAGTTAGAGCAATT 668
Query 1197 AAGCTTGCCCATGTGCAATGTGACATATGCAAAAGCTTGGGAATGAGGTGAGT 1254

Db 669 GTACCTGTCTATGTGCAAACTTATATATGAAAGCCCTGGAGATGTGCGCAATTT 726

RESULT 5
US-09-090-567-3
Sequence 3, Application US/09090567
Patent No. 5989549
GENERAL INFORMATION:
APPLICANT: Sullivan, Robert
APPLICANT: Bull, Bruno
APPLICANT: Igar, Christine
TITLE OF INVENTION: Acrosomal Sperm Protein And
TITLE OF INVENTION: Uses Thereof
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSER: Swabey Ogilvy Renault
STREET: 1600 - 1981 McGill College
CITY: Montreal
STATE: QC
COUNTRY: Canada
ZIP: H3A 2Y3
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/090,567
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Murphy, Kevin P
REGISTRATION NUMBER: 26,674
REFERENCE/DOCKET NUMBER: 13045-2"US" FC/CC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 514-845-7126
TELEFAX: 514-288-8389
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 912 base pairs
TYPE: nucleic acid
STRANDS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
US-09-090-567-3
Query Match 4.3%; Score 77; DB 2; Length 912;
Best Local Similarity 94.1%; Pred. No. 1.8e-11;
Matches 80; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Query 1 AGCTGAGCTCCACCGCGGCGGCGGCGGCTCTAAGACCTAGTGAATCCCGGCGTGGAG 60
Db 7 AGCTGAGCTCCACCGCGGCGGCGGCGGCTCTAAGACCTAGTGAATCCCGGCGTGGAG 66
Query 61 AATTGGCAGAGCCGAGAGCTGC 85
Db 67 AATTGGCAGAGCCGAGAGCTGC 91
RESULT 6
US-08-680-395-4
Sequence 4, Application US/08680395
Patent No. 5892010
GENERAL INFORMATION:
APPLICANT: Gray, Joe W.
APPLICANT: Collins, Colin

```

; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 2563
; TYPE: DNA
; ORGANISM: Glycine max
US-09-857-556A-9

Query Match
Best Local Similarity 87.4%; Score 75.6; DB 4; Length 2583;
Matches 83; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 AGCTGAAGCTTCACCCGGTGGCGCCGCTCTAGAACTAGTGGATTCGCCGGGCTGAGG 60
DB 46 AGCTGAAGCTTCACCCGGTGGCGCCGCTCTAGAACTAGTGGATTCGCCGGGCTGAGG 105
QY 61 AATTCGACGACGACGAGAGAGTGCATCTGACGAG 95
DB 106 AATTCGACGACGAGGTGATTTACTTTGTTCTCA 140

RESULT 9
US-09-904-615-32
; Sequence 32, Application US/09904615
; Patent No. 6566325
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: P2032P1
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32
; LENGTH: 1906
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: SITE
; LOCATION: (617)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (940)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1461)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1801)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-904-615-32

Query Match
Best Local Similarity 95.1%; Score 75.6; DB 4; Length 1906;
Matches 78; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CTGGAAGCTTCACCCGGTGGCGCCGCTCTAGAACTAGTGGATTCGCCGGGCTGAGG 62
DB 2 CTGGAAGCTTCACCCGGTGGCGCCGCTCTAGAACTAGTGGATTCGCCGGGCTGAGG 61
QY 63 TTGCGACGACGACGAGAGCTG 84
DB 62 TTGCGACGACGACGAGCTTGTG 83

RESULT 9
US-09-268-311-17
```

```
/ Sequence 17, Application US/09268311
/ Patent No. 6482923
/ GENERAL INFORMATION:
/ APPLICANT: Shi, Yangu
/ APPLICANT: Ruben, Steve M.
/ TITLE OF INVENTION: Interleukin 17 Receptor-Like Protein
/ FILE REFERENCE: PF398PI
/ CURRENT APPLICATION NUMBER: US/09/268,311
/ EARLIER FILING DATE: 1999-03-16
/ EARLIER APPLICATION NUMBER: 60/059,133
/ EARLIER FILING DATE: 1997-09-17
/ EARLIER APPLICATION NUMBER: 09/154,219
/ EARLIER FILING DATE: 1998-09-16
/ NUMBER OF SEQ ID NOS: 18
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 17
/ LENGTH: 1918
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (111)..(1409)
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (29)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (101)
/ OTHER INFORMATION: n equals a, t, g or c
/ US-09-268-311-17

Query Match      4.3%; Score 75.6; DB 4; Length 1918;
Best Local Similarity 89.0%; Pred. No. 6.8e-11;
Matches 81; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy      1 AGCTGAGCTCCACGCGGCTGTGAGCTGATCCCGGCTGACAG 60
Db      36 AGCTGAGCTCCACGCGGCTGTGAGCTGATCCCGGCTGACAG 95

Qy      61 AATTGGCAGCAGCGGCTGATCCCGGCTGATCCCGGCTGATCC 91
Db      96 AATTGGCAGCAGCGGCTGATCCCGGCTGATCCCGGCTGATCC 126

RESULT 10
US-09-501-115-3
/ Sequence 3, Application US/09501115
/ Patent No. 6552249
/ GENERAL INFORMATION:
/ APPLICANT: Cahoon, Rebecca E.
/ APPLICANT: Fader, Gary M.
/ APPLICANT: Rafaleki, Antoni
/ TITLE OF INVENTION: Plant Glnamyl-Alcohol Dehydrogenase Homologs
/ FILE REFERENCE: BM1328 US NA
/ CURRENT APPLICATION NUMBER: US/09/501,115
/ CURRENT FILING DATE: 2000-02-09
/ EARLIER APPLICATION NUMBER: 60/119,585
/ EARLIER FILING DATE: 1999-February-10
/ NUMBER OF SEQ ID NOS: 48
/ SOFTWARE: Microsoft Office 97
/ SEQ ID NO 3
/ LENGTH: 1463
/ TYPE: DNA
/ ORGANISM: Glycine max
/ US-09-501-115-3

Query Match      4.2%; Score 74.6; DB 4; Length 1463;
Best Local Similarity 95.1%; Pred. No. 1.1e-10;
Matches 77; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      1 AGCTGAGCTCCACGCGGCTGTGAGCTGATCCCGGCTGATCC 60
Db      1 AGCTGAGCTCCACGCGGCTGTGAGCTGATCCCGGCTGATCC 60
```

```
Db      41 AGCTGAGCTCCACGCGGCTGTGAGCTGATCCCGGCTGATCC 100
Qy      61 AATTGGCAGCAGCGGCTGATCCCGGCTGATCCCGGCTGATCC 81
Db      101 AATTGGCAGCAGCGGCTGATCCCGGCTGATCCCGGCTGATCC 121

RESULT 11
US-09-489-847-17
/ Sequence 17, Application US/09489847
/ Patent No. 6476195
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al
/ TITLE OF INVENTION: 98 Human Secreted Proteins
/ FILE REFERENCE: P2031PI
/ CURRENT APPLICATION NUMBER: US/09/489,847
/ CURRENT FILING DATE: 2000-01-24
/ EARLIER APPLICATION NUMBER: PCT/US99/17130
/ EARLIER FILING DATE: 1999-07-29
/ EARLIER APPLICATION NUMBER: 60/094,657
/ EARLIER FILING DATE: 1998-07-30
/ EARLIER APPLICATION NUMBER: 60/095,486
/ EARLIER FILING DATE: 1998-08-05
/ EARLIER APPLICATION NUMBER: 60/096,319
/ EARLIER FILING DATE: 1998-08-12
/ EARLIER APPLICATION NUMBER: 60/095,454
/ EARLIER FILING DATE: 1998-08-06
/ EARLIER APPLICATION NUMBER: 60/095,455
/ EARLIER FILING DATE: 1998-08-06
/ NUMBER OF SEQ ID NOS: 376
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 17
/ LENGTH: 1774
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ US-09-489-847-17

Query Match      4.2%; Score 74.6; DB 4; Length 1774;
Best Local Similarity 95.1%; Pred. No. 1.2e-10;
Matches 77; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      1 AGCTGAGCTCCACGCGGCTGTGAGCTGATCCCGGCTGATCC 60
Db      141 AGCTGAGCTCCACGCGGCTGTGAGCTGATCCCGGCTGATCC 200

Qy      61 AATTGGCAGCAGCGGCTGATCCCGGCTGATCCCGGCTGATCC 81
Db      201 AATTGGCAGCAGCGGCTGATCCCGGCTGATCCCGGCTGATCC 221

RESULT 12
US-09-489-847-56
/ Sequence 56, Application US/09489847
/ Patent No. 6476195
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al
/ TITLE OF INVENTION: 98 Human Secreted Proteins
/ FILE REFERENCE: P2031PI
/ CURRENT APPLICATION NUMBER: US/09/489,847
/ CURRENT FILING DATE: 2000-01-24
/ EARLIER APPLICATION NUMBER: PCT/US99/17130
/ EARLIER FILING DATE: 1999-07-29
/ EARLIER APPLICATION NUMBER: 60/094,657
/ EARLIER FILING DATE: 1998-07-30
/ EARLIER APPLICATION NUMBER: 60/095,486
/ EARLIER FILING DATE: 1998-08-05
/ EARLIER APPLICATION NUMBER: 60/096,319
/ EARLIER FILING DATE: 1998-08-12
/ EARLIER APPLICATION NUMBER: 60/095,454
/ EARLIER FILING DATE: 1998-08-06
/ EARLIER APPLICATION NUMBER: 60/095,455
/ EARLIER FILING DATE: 1998-08-06
/ NUMBER OF SEQ ID NOS: 376
```

```
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 56
/ LENGTH: 632
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURES:
/ NAME/KEY: SITE
/ LOCATION: (29)
/ OTHER INFORMATION: n equals a,t,g, or c
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (46)
/ OTHER INFORMATION: n equals a,t,g, or c
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (94)
/ OTHER INFORMATION: n equals a,t,g, or c
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (162)
/ OTHER INFORMATION: n equals a,t,g, or c
US-09-489-847-56

Query Match
Best Local Similarity 96.2%; Score 74.2; DB 4; Length 632;
Best Local Similarity 96.2%; Pred. No. 9e-11;
Matches 76; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AGCTGAGCTCCACCGCGGCTGAGACTAGTGATCCCGCGGCTGCAGG 60
DB 167 AGCTGAGCTCCACCGCGGCTGAGACTAGTGATCCCGCGGCTGCAGG 225
QY 61 AATTGGCAGCAGCGGAGG 79
DB 227 AATTGGCAGCAGCGGAGG 245

RESULT 13
US-09-489-847-62
/ Sequence 62, Application US/09489847
/ Patent No. 6476195
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al
/ TITLE OF INVENTION: 98 Human Secreted Proteins
/ FILE REFERENCE: P2031P1
/ CURRENT APPLICATION NUMBER: US/09/489,847
/ EARLIER FILING DATE: 2000-01-24
/ EARLIER APPLICATION NUMBER: PCT/US99/11130
/ EARLIER FILING DATE: 1999-07-29
/ EARLIER APPLICATION NUMBER: 60/094,657
/ EARLIER FILING DATE: 1998-07-30
/ EARLIER APPLICATION NUMBER: 60/095,486
/ EARLIER FILING DATE: 1998-08-05
/ EARLIER APPLICATION NUMBER: 60/096,319
/ EARLIER FILING DATE: 1998-08-12
/ EARLIER APPLICATION NUMBER: 60/095,454
/ EARLIER FILING DATE: 1998-08-06
/ EARLIER APPLICATION NUMBER: 60/095,455
/ EARLIER FILING DATE: 1998-08-06
/ NUMBER OF SEQ ID NOS: 376
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 62
/ LENGTH: 762
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURES:
/ NAME/KEY: SITE
/ LOCATION: (10)
/ OTHER INFORMATION: n equals a,t,g, or c
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (12)
/ OTHER INFORMATION: n equals a,t,g, or c
/ FEATURE:
```

```
/ NAME/KEY: SITE
/ LOCATION: (42)
/ OTHER INFORMATION: n equals a,t,g, or c
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (219)
/ OTHER INFORMATION: n equals a,t,g, or c
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (747)
/ OTHER INFORMATION: n equals a,t,g, or c
US-09-489-847-62

Query Match
Best Local Similarity 93.9%; Score 74; DB 4; Length 762;
Best Local Similarity 93.9%; Pred. No. 1.1e-10;
Matches 77; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AGCTGAGCTCCACCGCGGCTGAGACTAGTGATCCCGCGGCTGCAGG 60
DB 293 AGCTGAGCTCCACCGCGGCTGAGACTAGTGATCCCGCGGCTGCAGG 352
QY 61 AATTGGCAGCAGCGGAGG 82
DB 353 AATTGGCAGCAGCGGAGG 374

RESULT 14
US-08-075-533-20
/ Sequence 20, Application US/08075533
/ Patent No. 5530186
/ GENERAL INFORMATION:
/ APPLICANT: Hitz, William D.
/ APPLICANT: Yadav, Narendra S.
/ TITLE OF INVENTION: Nucleotide Sequences of Soybean Acyl-ACP
/ NUMBER OF SEQUENCES: 22
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: E. I. du Pont de Nemours and Company
/ STREET: 1007 Market Street
/ CITY: Wilmington
/ STATE: Delaware
/ COUNTRY: U.S.A.
/ ZIP: 19898
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/075,533
/ FILING DATE:
/ CLASSIFICATION: 800
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/631,264
/ FILING DATE: 20-DEC-1990
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Morrissey, Bruce W.
/ REGISTRATION NUMBER: 30,663
/ REFERENCE/DOCKET NUMBER: CR-8926-A
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (302) 992-4927
/ TELEFAX: (302) 892-7949
/ TELEX: 835420
/ INFORMATION FOR SRO ID NO: 20:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1378 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ ORIGINAL SOURCE:
/ ORGANISM: Brassica napus
US-08-075-533-20
```


Query Match

4.2%; Score 73.8; DB 1; Length 1378;

Best Local Similarity 97.4%; Pred. No. 1.8e-10;

Matches 75; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY

1 AGCTGAGCTCCACCGGCTGCGGCGCTCTAGAACTAGTGTATCCCCCGGCTGCAAG 60

DB

52 AGCTGAGCTCCACCGGCTGCGGCGCTCTAGAACTAGTGTATCCCCCGGCTGCAAG 111

QY

61 AATGCGCAGAGCCGA 77

DB

112 AATGCGCAGAGAGGA 128

RESULT 15

US-08-948-176-20

; Sequence 20, Application US/08948176

; Patent No. 5945585

; GENERAL INFORMATION:

; APPLICANT: HITZ, WILLIAM D.

; TITLE OF INVENTION: ACYL-ACP THIOESTERASES GENES

; TITLE OF INVENTION: AND THEIR USE IN ALTERING PLANT

; TITLE OF INVENTION: OIL COMPOSITION

; NUMBER OF SEQUENCES: 26

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY

; STREET: 1007 MARKET STREET

; CITY: WILMINGTON

; STATE: DELAWARE

; COUNTRY: UNITED STATES OF AMERICA

; ZIP: 19898

; COMPUTER READABLE FORM:

; MEDIUM TYPE: DISKETTE, 3.50 INCH

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95

; SOFTWARE: MICROSOFT WORD VERSION 7.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/948,176

; FILING DATE:

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/631,264

; FILING DATE: DECEMBER 20, 1990

; ATTORNEY/AGENT INFORMATION:

; NAME: CRISTENBURY, LYNN M.

; REGISTRATION NUMBER: 30,971

; REFERENCE/DOCKET NUMBER: CR-8926-C

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 302-992-5481

; TELEFAX: 302-773-0164

; INFORMATION FOR SEQ ID NO: 20:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1378 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; ORIGINAL SOURCE:

; ORGANISM: Brassica napus

; US-08-948-176-20

Query Match

4.2%; Score 73.8; DB 2; Length 1378;

Best Local Similarity 97.4%; Pred. No. 1.8e-10;

Matches 75; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY

1 AGCTGAGCTCCACCGGCTGCGGCGCTCTAGAACTAGTGTATCCCCCGGCTGCAAG 60

DB

52 AGCTGAGCTCCACCGGCTGCGGCGCTCTAGAACTAGTGTATCCCCCGGCTGCAAG 111

QY

61 AATGCGCAGAGCCGA 77

DB

112 AATGCGCAGAGAGGA 128

Search completed: October 6, 2003, 14:43:49

Job time: 122 secs